ELEMENTS OF PHOTOGRAMMETRY WITH APPLICATIONS IN GIS

FOURTH EDITION

PAUL R. WOLF BON A. DEWITT BENJAMIN E. WILKINSON

Elements of Photogrammetry with Applications in GIS

. Paul R. Wolf, Ph.D
Bon A. Dewitt, Ph.D.
.Benjamin E. Wilkinson, Ph.D

Fourth Edition

New York Chicago San Francisco
Athens London Madrid
Mexico City Milan New Delhi
Singapore Sydney Toronto

Copyright © 2014 by McGraw-Hill Education. All rights reserved. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

ISBN: 978-0-07-176111-6

MHID: 0-07-176111-X

e-Book conversion by Cenveo® Publisher Services

Version 1.0

The material in this eBook also appears in the print version of this title: ISBN: 978-0-07-176112-3, MHID: 0-07-176112-8.

McGraw-Hill Education eBooks are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. To contact a representative, please visit the Contact Us page at www.mhprofessional.com.

All trademarks are trademarks of their respective owners. Rather than put a trademark symbol after every occurrence of a trademarked name, we use names in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark. Where such designations appear in this book, they have been printed with initial caps.

Information has been obtained by McGraw-Hill Education from sources believed to be reliable. However, because of the possibility of human or mechanical error by our sources, McGraw-Hill Education, or others, McGraw-Hill Education does not guarantee the accuracy, adequacy, or completeness of any information and is not responsible for any errors or omissions or the results obtained from the use of such information.

TERMS OF USE

This is a copyrighted work and McGraw-Hill Education and its licensors reserve all rights in and to the work. Use of this work is subject to these terms. Except as permitted under the Copyright Act of 1976 and the right to store and retrieve one copy of the work, you may not decompile, disassemble, reverse engineer, reproduce, modify, create derivative works based upon, transmit, distribute, disseminate, sell, publish or sublicense the work or any part of it without McGraw-Hill Education's prior consent. You may use the work for your own noncommercial and personal use; any other use of the work is strictly prohibited. Your right to use the work may be terminated if you fail to comply with these terms.

THE WORK IS PROVIDED "AS IS." McGRAW-HILL EDUCATION AND ITS LICENSORS MAKE NO GUARANTEES OR WARRANTIES AS TO THE ACCURACY, ADEQUACY OR COMPLETENESS OF OR RESULTS TO BE OBTAINED FROM USING THE WORK, INCLUDING ANY INFORMATION THAT CAN BE ACCESSED THROUGH THE WORK VIA HYPERLINK OR OTHERWISE, AND EXPRESSLY DISCLAIM ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. McGraw-Hill Education and its licensors do not warrant or guarantee that the functions contained in the work will meet your requirements or that its operation will be uninterrupted or error free. Neither McGraw-Hill Education nor its licensors shall be liable to you or anyone else for any inaccuracy, error or omission, regardless of cause, in the work or for any damages resulting therefrom. McGraw-Hill Education has no responsibility for the content of any information accessed through the work. Under no circumstances shall McGraw-Hill Education and/or its licensors be liable for any indirect, incidental, special, punitive, consequential or similar damages that result from the use of or inability to use the work, even if any of them has been advised of the possibility of such damages. This limitation of liability shall apply to any claim or cause whatsoever whether such claim or cause arises in contract, tort or otherwise.

Contents

1 Introduction
- 1-1 Definition of Photogrammetry
- 1-2 History of Photogrammetry
- 1-3 Types of Photographs
- 1-4 Taking Vertical Aerial Photographs
- 1-5 Existing Aerial Photography
- 1-6 Uses of Photogrammetry
- 1-7 Photogrammetry and Geographic Information Systems
- 1-8 Professional Photogrammetry Organizations

References
Problems

2 Principles of Photography and Imaging
- 2-1 Introduction
- 2-2 Fundamental Optics
- 2-3 Lenses
- 2-4 Single-Lens Camera
- 2-5 Illuminance
- 2-6 Relationship of Aperture and Shutter Speed
- 2-7 Characteristics of Photographic Emulsions
- 2-8 Processing and Printing Black-and-White Photographs
- 2-9 Spectral Sensitivity of Emulsions
- 2-10 Filters
- 2-11 Color Film
- 2-12 Digital Images
- 2-13 Color Image Representation
- 2-14 Digital Image Display

References
Problems

3 Cameras and Other Imaging Devices
- 3-1 Introduction
- 3-2 Metric Cameras for Aerial Mapping
- 3-3 Main Parts of Frame Aerial Cameras
- 3-4 Focal Plane and Fiducial Marks
- 3-5 Shutters
- 3-6 Camera Mounts
- 3-7 Camera Controls
- 3-8 Automatic Data Recording
- 3-9 Digital Mapping Cameras
- 3-10 Camera Calibration
- 3-11 Laboratory Methods of Camera Calibration
- 3-12 Stellar and Field Methods of Camera Calibration

3-13 Calibration of Nonmetric Cameras
3-14 Calibrating the Resolution of a Camera
References
Problems

4 Image Measurements and Refinements
4-1 Introduction
4-2 Coordinate Systems for Image Measurements
4-3 Simple Scales for Photographic Measurements
4-4 Measuring Photo Coordinates with Simple Scales
4-5 Comparator Measurement of Photo Coordinates
4-6 Photogrammetric Scanners
4-7 Refinement of Measured Image Coordinates
4-8 Distortions of Photographic Films and Papers
4-9 Image Plane Distortion
4-10 Reduction of Coordinates to an Origin at the Principal Point
4-11 Correction for Lens Distortions
4-12 Correction for Atmospheric Refraction
4-13 Correction for Earth Curvature
4-14 Measurement of Feature Positions and Edges
References
Problems

5 Object Space Coordinate Systems
5-1 Introduction
5-2 Concepts of Geodesy
5-3 Geodetic Coordinate System
5-4 Geocentric Coordinates
5-5 Local Vertical Coordinates
5-6 Map Projections
5-7 Horizontal and Vertical Datums
References
Problems

6 Vertical Photographs
6-1 Geometry of Vertical Photographs
6-2 Scale
6-3 Scale of a Vertical Photograph Over Flat Terrain
6-4 Scale of a Vertical Photograph Over Variable Terrain
6-5 Average Photo Scale
6-6 Other Methods of Determining Scale of Vertical Photographs
6-7 Ground Coordinates from a Vertical Photograph
6-8 Relief Displacement on a Vertical Photograph
6-9 Flying Height of a Vertical Photograph
6-10 Error Evaluation
References
Problems

7 Stereoscopic Viewing
7-1 Depth Perception
7-2 The Human Eye
7-3 Stereoscopic Depth Perception
7-4 Viewing Photographs Stereoscopically
7-5 Stereoscopes
7-6 The Use of Stereoscopes
7-7 Causes of *Y* Parallax
7-8 Vertical Exaggeration in Stereoviewing
References
Problems

8 Stereoscopic Parallax
8-1 Introduction
8-2 Photographic Flight-Line Axes for Parallax Measurement
8-3 Monoscopic Methods of Parallax Measurement
8-4 Principle of the Floating Mark
8-5 Stereoscopic Methods of Parallax Measurement
8-6 Parallax Equations
8-7 Elevations by Parallax Differences
8-8 Simplified Equation for Heights of Objects from Parallax Differences
8-9 Measurement of Parallax Differences
8-10 Computing Flying Height and Air Base
8-11 Error Evaluation
References
Problems

9 Elementary Methods of Planimetric Mapping for GIS
9-1 Introduction
9-2 Planimetric Mapping with Reflection Instruments
9-3 Georeferencing of Digital Imagery
9-4 Heads-Up Digitizing
9-5 Photomaps
9-6 Mosaics
9-7 Uncontrolled Digital Mosaics
9-8 Semicontrolled Digital Mosaics
9-9 Controlled Digital Mosaics
References
Problems

10 Tilted and Oblique Photographs
10-1 Introduction
10-2 Point Perspective
10-3 Angular Orientation in Tilt, Swing, and Azimuth
10-4 Auxiliary Tilted Photo Coordinate System
10-5 Scale of a Tilted Photograph
10-6 Relief Displacement on a Tilted Photograph

10-7 Determining the Angle of Inclination of the Camera Axis in Oblique Photography
10-8 Computing Horizontal and Vertical Angles from Oblique Photos
10-9 Angular Orientation in Omega-Phi-Kappa
10-10 Determining the Elements of Exterior Orientation
10-11 Rectification of Tilted Photographs
10-12 Correction for Relief of Ground Control Points Used in Rectification
10-13 Analytical Rectification
10-14 Optical-Mechanical Rectification
10-15 Digital Rectification
10-16 Atmospheric Refraction in Tilted Aerial Photographs
References
Problems

11 Introduction to Analytical Photogrammetry
11-1 Introduction
11-2 Image Measurements
11-3 Control Points
11-4 Collinearity Condition
11-5 Coplanarity Condition
11-6 Space Resection by Collinearity
11-7 Space Intersection by Collinearity
11-8 Analytical Stereomodel
11-9 Analytical Interior Orientation
11-10 Analytical Relative Orientation
11-11 Analytical Absolute Orientation
References
Problems

12 Stereoscopic Plotting Instruments
12-1 Introduction
12-2 Classification of Stereoscopic Plotters

PART I DIRECT OPTICAL PROJECTION STEREOPLOTTERS

12-3 Components
12-4 Projection Systems
12-5 Viewing and Tracing Systems
12-6 Interior Orientation
12-7 Relative Orientation
12-8 Absolute Orientation

PART II ANALYTICAL PLOTTERS

12-9 Introduction
12-10 System Components and Method of Operation
12-11 Analytical Plotter Orientation
12-12 Three-Dimensional Operation of Analytical Plotters
12-13 Modes of Use of Analytical Plotters

PART III SOFTCOPY PLOTTERS

12-14 Introduction
12-15 System Hardware
12-16 Image Measurements
12-17 Orientation Procedures
12-18 Epipolar Geometry
References
Problems

13 Topographic Mapping and Spatial Data Collection
13-1 Introduction
13-2 Direct Compilation of Planimetric Features by Stereoplotter
13-3 Direct Compilation of Contours by Stereoplotter
13-4 Digitizing Planimetric Features from Stereomodels
13-5 Representing Topographic Features in Digital Mapping
13-6 Digital Elevation Models and Indirect Contouring
13-7 Automatic Production of Digital Elevation Models
13-8 Orthophoto Generation
13-9 Map Editing
References
Problems

14 Laser Scanning Systems
14-1 Introduction
14-2 Principles and Hardware
14-3 Airborne Laser Scanning
14-4 Terrestrial Laser Scanning
14-5 Laser Scan Data
14-5 Error Evaluation
References
Problems

15 Fundamental Principles of Digital Image Processing
15-1 Introduction
15-2 The Digital Image Model
15-3 Spatial Frequency of a Digital Image
15-4 Contrast Enhancement
15-5 Spectral Transformations
15-6 Moving Window Operations
15-7 Multiscale Representation
15-8 Digital Image Matching
15-9 Summary
References
Problems

16 Control for Aerial Photogrammetry
16-1 Introduction

16-2 Ground Control Images and Artificial Targets
16-3 Number and Location of Photo Control
16-4 Traditional Field Survey Methods for Establishing Horizontal and Vertical Control
16-5 Fundamentals of the Global Positioning System
16-6 Kinematic GPS Positioning
16-7 Inertial Navigation Systems
16-8 GPS-INS Integration
References
Problems

17 Aerotriangulation
17-1 Introduction
17-2 Pass Points for Aerotriangulation
17-3 Fundamentals of Semianalytical Aerotriangulation
17-4 Sequential Construction of a Strip Model from Independent Models
17-5 Adjustment of a Strip Model to Ground
17-6 Simultaneous Bundle Adjustment
17-7 Initial Approximations for the Bundle Adjustment
17-8 Bundle Adjustment with Airborne GPS Control
17-9 Interpretation of Bundle Adjustment Results
17-10 Aerotriangulation with Airborne Linear Array Sensors
17-11 Satellite Image Triangulation
17-12 Efficient Computational Strategies for Aerotriangulation
References
Problems

18 Project Planning
18-1 Introduction
18-2 Importance of Flight Planning
18-3 Photographic End Lap and Side Lap
18-4 Purpose of the Photography
18-5 Photo Scale
18-6 Flying Height
18-7 Ground Coverage
18-8 Weather Conditions
18-9 Season of the Year
18-10 Flight Map
18-11 Specifications
18-12 Cost Estimating and Scheduling
References
Problems

19 Terrestrial and Close-Range Photogrammetry
19-1 Introduction
19-2 Applications of Terrestrial and Close-Range Photogrammetry
19-3 Terrestrial Cameras
19-4 Matrix Equations for Analytical Self-Calibration

19-5 Initial Approximations for Least Squares Adjustment
19-6 Solution Approach for Self-Calibration Adjustment
19-7 Control for Terrestrial Photogrammetry
19-8 Analytical Self-Calibration Example
19-9 Planning for Close-Range Photogrammetry
References
Problems

20 Photogrammetric Applications in GIS
20-1 Introduction
20-2 Land and Property Management
20-3 Floodplain Rating
20-4 Water Quality Management
20-5 Wildlife Management
20-6 Environmental Monitoring
20-7 Wetland Analysis
20-8 Transportation
20-9 Multipurpose Land Information System
20-10 Summary
References
Problems

A Units, Errors, Significant Figures, and Error Propagation

B Introduction to Least Squares Adjustment

C Coordinate Transformations

D Development of Collinearity Condition Equations

E Digital Resampling

F Conversions Between Object Space Coordinate Systems

Index

About the Authors

Paul R. Wolf, Ph.D. (deceased), was a professor emeritus of civil and environmental engineering at the University of Wisconsin, Madison.

Bon A. Dewitt, Ph.D., is an associate professor in the Geomatics Program, School of Forest Resources and Conservation, University of Florida, Gainesville. Since 1999 he has served as the director of the program. Dr. Dewitt specializes in photogrammetry, digital mapping technology, digital image processing, hydrographic surveys, subdivision design, and land surveying.

Benjamin E. Wilkinson, Ph.D., is an assistant professor in the Geomatics Program, School of Forest Resources and Conservation, University of Florida, Gainesville. Previously, he was a research scientist at Integrity Applications Incorporated. Dr. Wilkinson specializes in photogrammetry, LiDAR, remote sensing, navigation, and software development.

CHAPTER 1
Introduction

1-1 Definition of Photogrammetry

Photogrammetry has been defined by the American Society for Photogrammetry and Remote Sensing as the art, science, and technology of obtaining reliable information about physical objects and the environment through processes of recording, measuring, and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena. As implied by its name, the science originally consisted of analyzing photographs, however the use of film cameras has greatly diminished in favor of digital sensors. Photogrammetry has expanded to include analysis of other records, such as digital imagery, radiated acoustical energy patterns, laser ranging measurements, and magnetic phenomena. In this text both *photographic* and *digital* photogrammetry are emphasized since they share many of the same principles, but other sources of information are also discussed. The terms *photograph* and *photo* as used in this book can be considered synonymous with *digital image* unless specifically noted.

Included within the definition of photogrammetry are two distinct areas: (1) *metric* photogrammetry and (2) *interpretative* photogrammetry. Metric photogrammetry consists of making precise measurements from photos and other information sources to determine, in general, the relative locations of points. This enables finding distances, angles, areas, volumes, elevations, and sizes and shapes of objects. The most common applications of metric photogrammetry are the preparation of planimetric and topographic maps from photographs (see Secs. 13-2 through 13-7), and the production of orthophotos from digital imagery (see Sec. 13-8). The photographs are most often *aerial* (taken from an airborne vehicle), but *terrestrial* photos (taken from earth-based cameras) and satellite imagery are also used.

Interpretative photogrammetry deals principally in recognizing and identifying objects and judging their significance through careful and systematic analysis. It is included in the branches of *image interpretation* and *remote sensing*. Image interpretation and remote sensing include not only the analysis of photography but also the use of data gathered from a wide variety of sensing instruments, including multispectral cameras, infrared sensors, thermal scanners, and sidelooking airborne radar. Remote sensing instruments, which are often carried in vehicles as remote as orbiting satellites, are capable of providing quantitative as well as qualitative information about objects. At present, with our recognition of the importance of preserving our environment and natural resources, photographic interpretation and remote sensing are both being employed extensively as tools in management and planning.

Of the two distinct areas of photogrammetry, concentration in this book is on metric photogrammetry. Interpretative photogrammetry is discussed only briefly, and those readers interested in further study in this area should consult the references cited at the end of this chapter.

1-2 History of Photogrammetry

Developments leading to the present-day science of photogrammetry occurred long before the invention of photography. As early as 350 B.C. Aristotle had referred to the process of projecting images optically. In the early 18th century Dr. Brook Taylor published his treatise on linear

perspective, and soon afterward, J. H. Lambert suggested that the principles of perspective could be used in preparing maps.

The actual practice of photogrammetry could not occur, of course, until a practical photographic process was developed. Pioneering research in this area was advanced by Joseph Niepce of France, who produced the world's first photograph in 1827 by a process he referred to as *heliography*. This process used metal plates coated with a tarlike substance that would gradually harden with exposure to light. Expanding on the work of Niepce, fellow Frenchman Louis Daguerre announced his direct photographic process, which was more practical than heliography. In his process the exposure was made on metal plates that had been light-sensitized with a coating of silver iodide. This is essentially the photographic process still in use today.

A year after Daguerre's announcement, Francois Arago, a geodesist with the French Academy of Science, demonstrated the use of photographs in topographic surveying. The first actual experiments in using photogrammetry for topographic mapping occurred in 1849 under the direction of Colonel Aimé Laussedat of the French Army Corps of Engineers. In Colonel Laussedat's experiments kites and balloons were used for taking aerial photographs. Due to difficulties encountered in obtaining aerial photographs, he curtailed this area of research and concentrated his efforts on mapping with terrestrial photographs. In 1859 Colonel Laussedat presented an account of his successes in mapping using photographs. His pioneering work and dedication to this subject earned him the title "father of photogrammetry."

Topographic mapping using photogrammetry was introduced to North America in 1886 by Captain Eduard Deville, the Surveyor General of Canada. He found Laussedat's principles extremely convenient for mapping the rugged mountains of western Canada. The U.S. Coast and Geodetic Survey (now the National Geodetic Survey) adopted photogrammetry in 1894 for mapping along the border between Canada and the Alaska Territory.

Meanwhile new developments in instrumentation, including improvements in cameras and films, continued to nurture the growth of photogrammetry. In 1861 a three-color photographic process was developed, and roll film was perfected in 1891. In 1909 Dr. Carl Pulfrich of Germany began to experiment with overlapping pairs of photographs. His work formed much of the foundation for the development of many instrumental photogrammetric mapping techniques in use today.

The invention of the airplane by the Wright brothers in 1903 provided the great impetus for the emergence of modern aerial photogrammetry. Until that time, almost all photogrammetric work was, for the lack of a practical means of obtaining aerial photos, limited to terrestrial photography. The airplane was first used in 1913 for obtaining photographs for mapping purposes. Aerial photos were used extensively during World War I, primarily in reconnaissance. In the period between World War I and World War II, aerial photogrammetry for topographic mapping progressed to the point of mass production of maps. Within this period many private firms and government agencies in North America and in Europe became engaged in photogrammetric work. During World War II, photogrammetric techniques were used extensively to meet the great new demand for maps. Air photo interpretation was also employed more widely than ever before in reconnaissance and intelligence. Out of this war-accelerated mapping program came many new developments in instruments and techniques.

Advancements in instrumentation and techniques in photogrammetry have continued at a rapid pace through the remainder of the 20th, and into the 21st century. The many advancements are too numerous to itemize here, but collectively they have enabled photogrammetry to become the most accurate and efficient method available for compiling maps and generating topographic information. The improvements have affected all aspects of the science, and they incorporate many new developments such as those in optics, electronics, computers and satellite technology. While this text

does include some historical background, its major thrust is to discuss and describe the current state of the art in photogrammetric instruments and techniques.

1-3 Types of Photographs

Two fundamental classifications of photography used in the science of photogrammetry are *terrestrial* and *aerial*. Terrestrial photographs (see Chap. 19) are taken with ground-based cameras, the position and orientation of which might be measured directly at the time of exposure. A great variety of cameras are used for taking terrestrial photographs, and these may include anything from inexpensive commercially available cameras to precise specially designed cameras. While there are still some film cameras being used in terrestrial photogrammetry, digital cameras have become the standard sensors for image acquisition.

Aerial photography is commonly classified as either vertical or oblique. *Vertical photos* are taken with the camera axis directed as nearly vertically as possible. If the camera axis were perfectly vertical when an exposure was made, the photographic plane would be parallel to the datum plane and the resulting photograph would be termed *truly vertical*. In practice, the camera axis is rarely held perfectly vertical due to unavoidable aircraft tilts. When the camera axis is unintentionally tilted slightly from vertical, the resulting photograph is called a *tilted photograph*. These unintentional tilts are usually less than 1° and seldom more than 3°. For many practical applications, simplified procedures suitable for analyzing truly vertical photos may also be used for tilted photos without serious consequence. Precise photogrammetric instruments and procedures have been developed, however, that make it possible to rigorously account for tilt with no loss of accuracy. Figure 1-1 shows a film-based aerial mapping camera with its electric control mechanism and the mounting framework for placing it in an aircraft. The vertical photograph illustrated in Fig. 1-2 was taken with a camera of the type illustrated in Fig. 1-1 from an altitude of 470 meters (m) above the terrain.

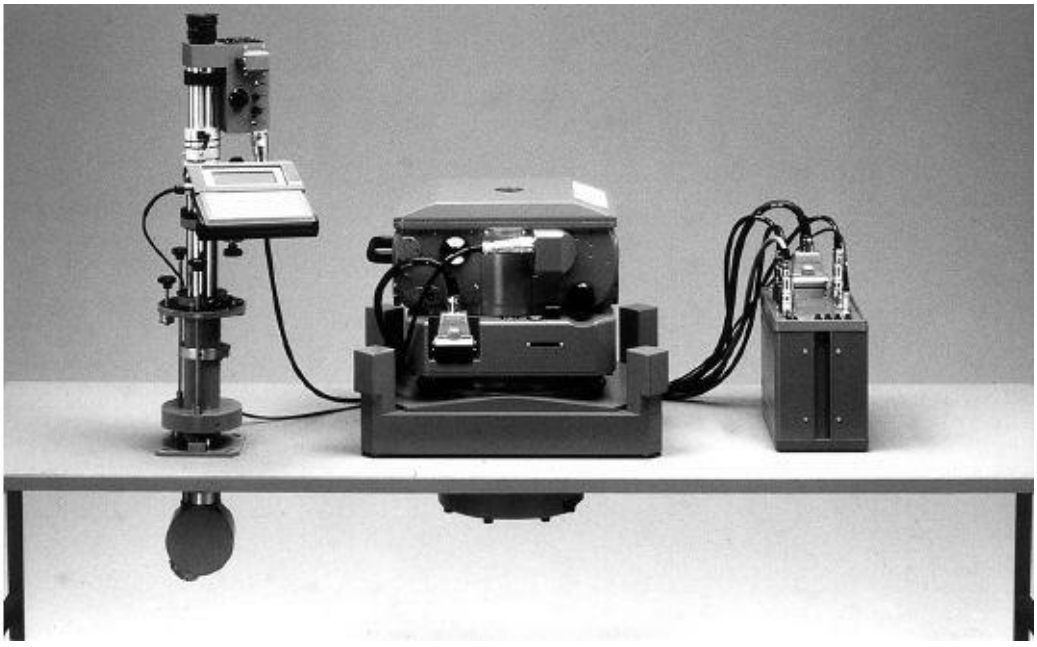


FIGURE 1-1 Zeiss RMK TOP 15, aerial mapping camera, with electronic controls and aircraft mountings. (*Courtesy Carl Zeiss, Inc.*)

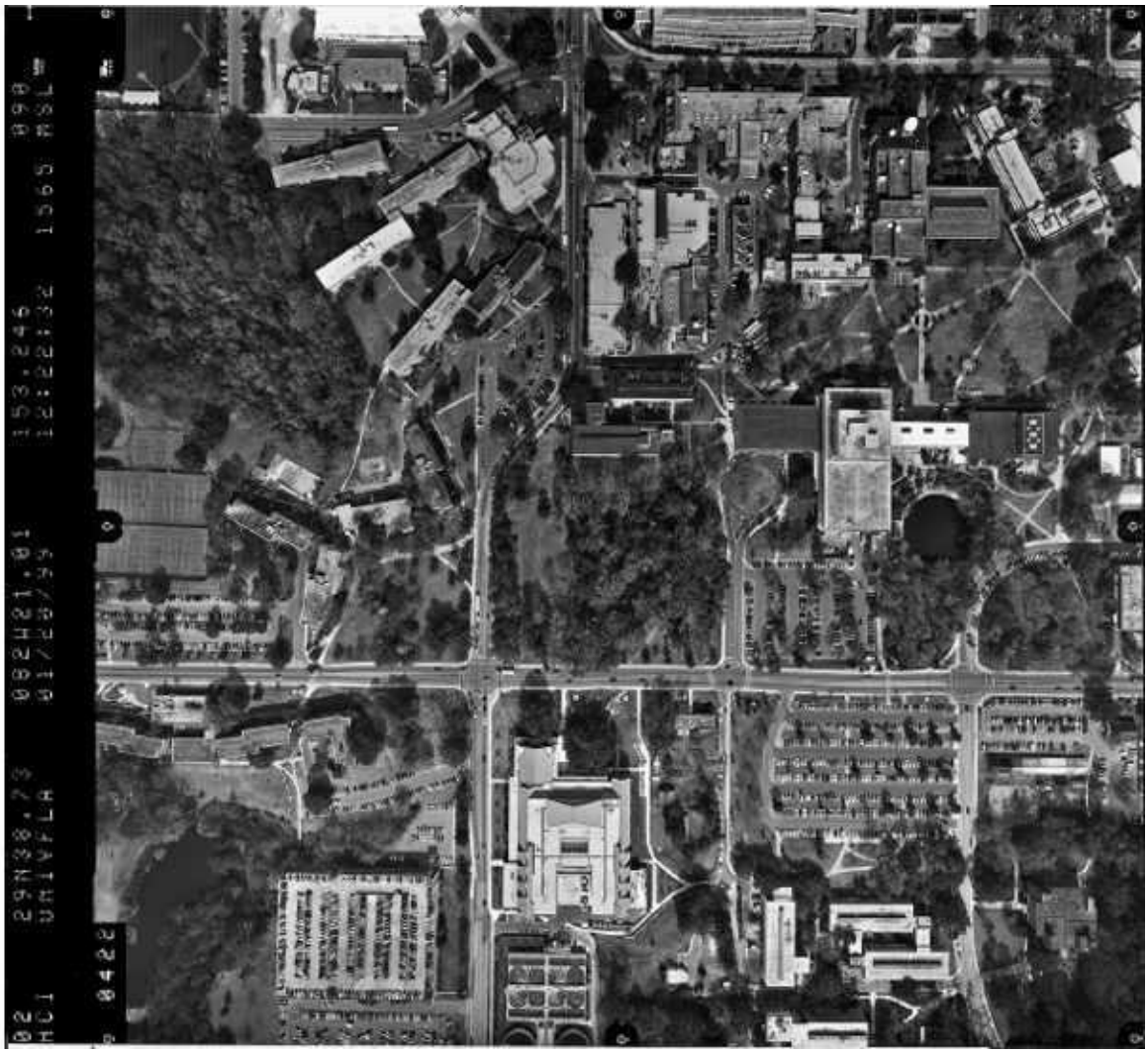


FIGURE 1-2 Vertical aerial photograph. (*Courtesy Hoffman and Company, Inc.*)

While numerous film-based aerial mapping cameras are still in use, they are steadily being replaced by high-resolution digital sensors. The sensor shown in Fig. 1-3 can capture digital images containing pictorial detail that rivals, and in some cases exceeds, that of film-based cameras. The geometry of the images produced by this sensor is effectively the same as that of standard film-based aerial mapping cameras, and thus allows the same analysis methods and equations. Figure 1-4 shows a digital sensor that acquires imagery by scanning the terrain continuously as the aircraft proceeds along its trajectory. This sensor requires special instrumentation that can determine the precise position and angular attitude as they vary continuously along the flight path. Substantial post-flight processing is required in order to produce undistorted images of the terrain from the raw data.

FIGURE 1-3 Microsoft UltraCam Eagle ultra-large digital aerial photogrammetric camera. (*Courtesy Microsoft Corporation.*)

FIGURE 1-4 Leica ADS80 airborne digital sensor. (*Courtesy Leica Geosystems.*)

Oblique aerial photographs are exposed with the camera axis intentionally tilted away from vertical. A *high oblique* photograph includes the horizon; a *low oblique* does not. Figure 1-5 illustrates the orientation of the camera for vertical, low oblique, and high oblique photography and also shows how a square grid of ground lines would appear in each of these types of photographs. Figures 1-6 and 1-7 are examples of low oblique and high oblique photographs, respectively.

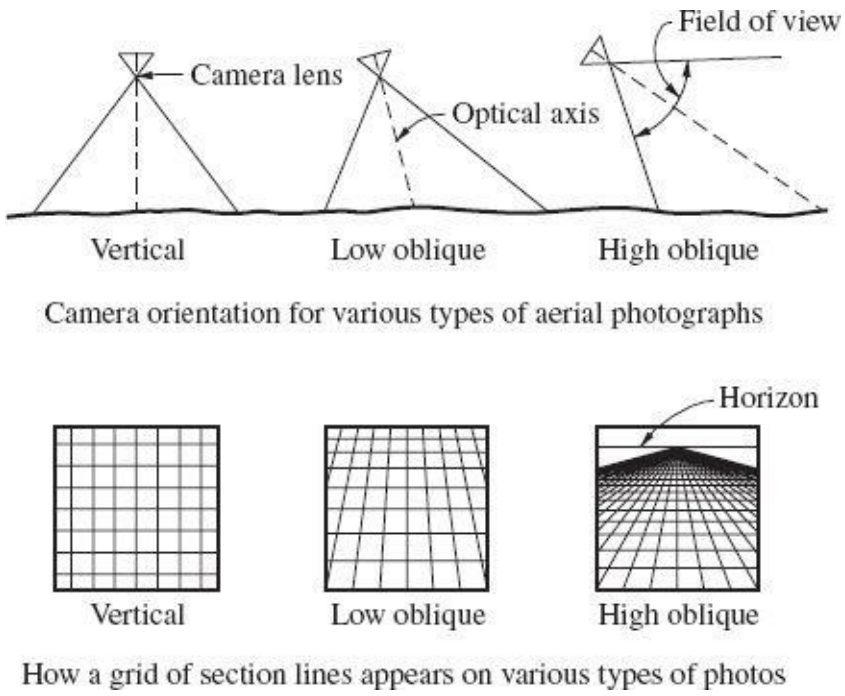


FIGURE 1-5 Camera orientation for various types of aerial photographs.

FIGURE 1-6 Low oblique photograph of Madison, Wisconsin (note that the horizon is not shown). (*Courtesy State of Wisconsin, Department of Transportation*.)

FIGURE 1-7 High oblique photograph of Tampa, Florida (note that the horizon shows on the photograph). (*Courtesy US Imaging, Inc.*)

Figure 1-8 is an example of a low oblique image taken with a digital camera. The camera's position and angular attitude was directly measured in order to precisely locate the image features in a ground coordinate system.

FIGURE 1-8 Low oblique digital camera image (*Courtesy Pictometry International Corp.*)

1-4 Taking Vertical Aerial Photographs

When an area is covered by vertical aerial photography, the photographs are usually taken along a series of parallel passes, called *flight strips*. As illustrated in Fig. 1-9, the photographs are normally exposed in such a way that the area covered by each successive photograph along a flight strip duplicates or overlaps part of the ground coverage of the previous photo. This lapping along the flight strip is called *end lap*, and the area of coverage common to an adjacent pair of photographs in a flight strip is called the *stereoscopic overlap area*. The overlapping pair of photos is called a *stereopair*. For reasons which will be given in subsequent chapters, the amount of end lap is normally between 55 and 65 percent. The positions of the camera at each exposure, e.g., positions 1, 2, 3 of Fig. 1-9, are called the *exposure stations*, and the .altitude of the camera at exposure time is called the *flying height*

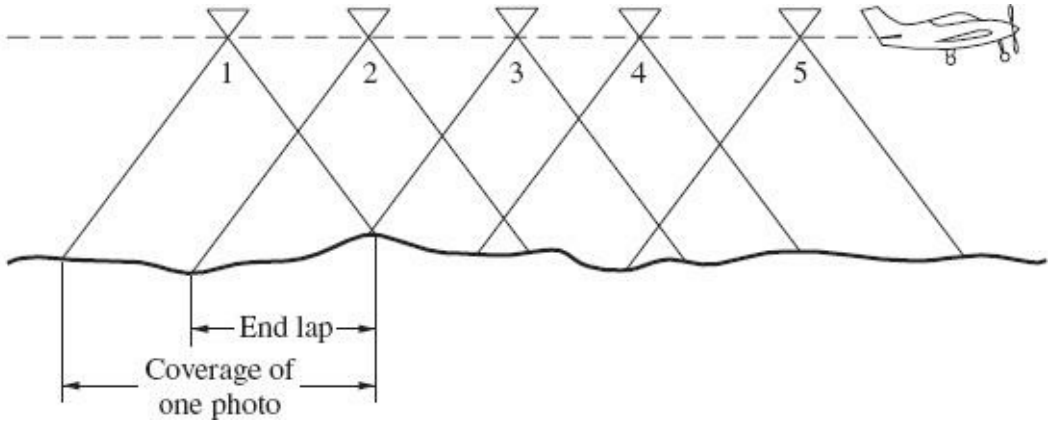


FIGURE 1-9 End lap of photographs in a flight strip.

Adjacent flight strips are photographed so that there is also a lateral overlapping of ground coverage between strips. This condition, as illustrated in Fig. 1-10, is called *side lap*, and it is normally held at approximately 30 percent. The photographs of two or more side-lapping strips used to cover an area is referred to as a *block* of photos.

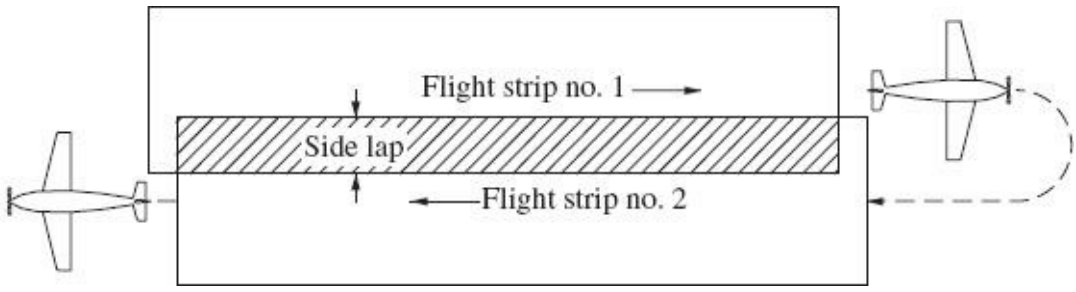


FIGURE 1-10 Side lap of adjacent flight strips.

1-5 Existing Aerial Photography

Photogrammetrists and photo interpreters can obtain aerial photography in one of two ways: (1) They can obtain photographs from existing coverage either for free or at a cost, or (2) they can purchase new coverage. It is seldom economical to use existing coverage for mapping, because it rarely meets the needs of the user; but existing coverage may prove suitable for other uses such as reconnaissance, project planning, historical records, or photo interpretation. If existing photography is not satisfactory because of age, scale, camera, etc., it will be necessary to obtain new coverage. Of course, before the decision can be made whether to use existing photography or obtain new, it is necessary to ascertain exactly what coverage exists in a particular area.

Existing aerial photography is available for nearly all the United States and Canada. Some areas have been covered several times, so that various scales and qualities of photography are available. Most of the coverage is vertical photography.

The U.S. Geological Survey (USGS), has archived, for distribution upon request,[1] millions of aerial photos and satellite images. These include air photo coverage of virtually all areas of the United States as well as images from several series of satellites which provide global coverage. Their archived air photo coverage includes photos taken through the *National Aerial Photography Program* (NAPP); they were taken from a flying height of 20,000 feet (ft) above ground and are in black and white and color-infrared. It also includes photos from the

National High Altitude Photography (NHAP) Program, also in black and white and color-infrared and taken from 40,000 ft above ground. The EROS Data Center also archives photos that were taken by the USGS for its topographic mapping projects as well as photos taken by other federal agencies including the *National Aeronautics and Space Administration* (NASA), the *Bureau of Reclamation*, the *Environmental Protection Agency* (EPA), and the *U.S. Army Corps of Engineers*.

The U.S. Department of Agriculture[2] is another useful resource for obtaining existing aerial photography. Their archives contain extensive coverage for the United States. Available products include black and white, color, and color infrared prints at negative scales of 1:20,000 and smaller.

Existing aerial photography can also be obtained from the department of transportation of most states. These photos have usually been taken for use in highway planning and design; thus the scales are generally relatively large, and coverage typically follows state and federal highways. In addition, many counties have periodic coverage.

1-6 Uses of Photogrammetry

The earliest applications of photogrammetry were in topographic mapping, and today that use is still the most common of photogrammetric activities. At present, the USGS, the federal agency charged with mapping the United States, performs nearly 100 percent of its map compilation photogrammetrically. State departments of transportation also use photogrammetry almost exclusively in preparing their topographic maps. In addition, private engineering and surveying firms prepare many special-purpose topographic maps photogrammetrically. These maps vary in scale from large to small and are used in planning and designing highways, railroads, rapid transit systems, bridges, pipelines, aqueducts, transmission lines, hydroelectric dams, flood control structures, river and harbor improvements, urban renewal projects, etc. A huge quantity of topographic maps are prepared for use in providing spatial data for geographic information systems (see Sec. 1-7).

Two newer photogrammetric products, *orthophotos* and *digital elevation models* (DEMs), are now often used in combination to replace traditional topographic maps. As described in Sec. 13-8, an orthophoto is an aerial photograph that has been modified so that its scale is uniform throughout. Thus orthophotos are equivalent to planimetric maps, but unlike planimetric maps which show features by means of lines and symbols, orthophotos show the actual images of features. For this reason they are more easily interpreted than planimetric maps, and hence are preferred by many users. A DEM, as discussed in Sec. 13-6, consists of an array of points in an area that have had their X, Y, and Z coordinates determined. Thus, they provide a numerical representation of the topography in the area, and contours, cross sections, profiles, etc., can be computed from them. Orthophotos and DEMs are both widely applied in all fields where maps are used, but because they are both in digital form, one of their most common applications, as discussed in Sec. 1-7 and Chap. 20, is their use in connection with geographic information systems.

Photogrammetry has become an exceptionally valuable tool in land surveying. To mention just a few uses in the field, aerial photos can be used as rough base maps for relocating existing property boundaries. If the point of beginning or any corners can be located with respect to ground features that can be identified on the photo, the entire parcel can be plotted on the photo from the property description. All corners can then be located on the photo in relation to identifiable ground features which, when located in the field, greatly assist in finding the actual property corners. Aerial photos can also be used in planning ground surveys. Through stereoscopic viewing, the area can be studied in three dimensions. Access routes to remote areas can be identified, and surveying lines of least resistance through difficult terrain or forests can be found. The photogrammetrist can prepare a map of an area without actually setting foot on the ground—an advantage which circumvents problems of

gaining access to private land for ground surveys.

The field of highway planning and design provides an excellent example of how important photogrammetry has become in engineering. In this field, high-altitude photos or satellite images are used to assist in area and corridor studies and to select the best route; smallscale topographic maps are prepared for use in preliminary planning; large-scale topographic maps and DEMs are compiled for use in final design; and earthwork cross sections are taken to obtain contract quantities. In many cases the plan portions of the "plan profile" sheets of highway plans are prepared from aerial photographs. Partial payments and even final pay quantities are often calculated from photogrammetric measurements. Map information collected from modern photogrammetric instruments is directly compatible with *computer-aided drafting* (CAD) systems commonly used in highway design. The use of photogrammetry in highway engineering not only has reduced costs but also has enabled better overall highway designs to be created.

Many areas outside of engineering have also benefitted from photogrammetry. Nonengineering applications include the preparation of tax maps, soil maps, forest maps, geologic maps, and maps for city and regional planning and zoning. Photogrammetry is used in the fields of astronomy, architecture, archaeology, geomorphology, oceanography, hydrology and water resources, conservation, ecology, and mineralogy. Stereoscopic photography effectively enables the outdoors to be brought into the comfortable confines of the laboratory or office for viewing and study in three dimensions.

Photogrammetry has been used successfully in traffic management and in traffic accident investigations. One advantage of its use in the latter area is that photographs overlook nothing that may be needed later to reconstruct the accident, and it is possible to restore normal traffic flow quickly. Even in the fields of medicine and dentistry, measurements from X-ray and other photographs and images have been useful in diagnosis and treatment. Of course, one of the oldest and still most important uses of aerial photography is in military intelligence. Space exploration is one of the new and exciting areas where photogrammetry is being utilized.

Photogrammetry has become a powerful research tool because it affords the unique advantage of permitting instantaneous recordings of dynamic occurrences to be captured in images. Measurements from photographs of quantities such as beam or pavement deflections under impact loads may easily be obtained photographically where such measurements would otherwise be nearly impossible.

As noted earlier, photogrammetry has become extremely important in the field of geographic information systems (GISs). This area of application is described in the following section.

It would be difficult to cover all the many situations in which photogrammetric principles and methods could be or have been used to solve measurement problems. Photogrammetry, although still a relatively new science, has already contributed substantially to engineering and nonengineering fields alike. New applications appear to be bounded only by our imagination, and the science should continue to grow in the future.

1-7 Photogrammetry and Geographic Information Systems

Geographic information systems, are widely used and hold a position of prominence in many fields. These computer-based systems enable storing, integrating, manipulating, analyzing, and displaying virtually any type of spatially related information about the environment. They are being used at all levels of government, and by businesses, private industry, and public utilities to assist in planning, design, management, and decision making. Chapter 20 in this book presents examples of applications in geographic information systems. It is important at this juncture, however, to mention the major role

that photogrammetry plays in these systems.

An essential element of any GIS is a complex relational database. The information that comprises the database usually includes both natural and cultural features. Specific types of information, *objects* o r *layers*, within the database may include political boundaries, individual property ownership, transportation networks, utilities, topography, hydrography, soil types, land use, vegetation types, wetlands, etc. To be of use in a GIS, however, all data must be spatially related; i.e., all the data must be in a common geographic frame of reference. Photogrammetry is ideal for deriving much of this spatial information. As noted in the preceding section, topographic maps, digital elevation models, and digital orthophotos are examples of photogrammetric products which are now commonly employed in developing these spatially related layers of information. By employing photogrammetry the data can be compiled more economically than through ground surveying methods, and this can be achieved with comparable or even greater spatial accuracy. Furthermore, the data are compiled directly in digital format, and thus are compatible for direct entry into GIS databases.

The photogrammetric procedures involved in developing topographic maps, digital elevation models, digital orthophotos, and other products used in GISs are described in later chapters of this text.

1-8 Professional Photogrammetry Organizations

There are numerous professional organizations in the United States and across the world serving the interests of photogrammetry. Generally these organizations have as their objectives the advancement of knowledge in the field, encouragement of communication among photogrammetrists, and upgrading of standards and ethics in the practice of photogrammetry.

The American Society for Photogrammetry and Remote Sensing (ASPRS), formerly known as the American Society of Photogrammetry, founded in 1934, is the foremost professional photogrammetric organization in the United States. One of this society's most valuable contributions has been its publication of various manuals, such as the *Manual of Photogrammetry*, the *Manual of Remote Sensing*, the *Manual of Photographic Interpretation*, and the *Manual of Geographic Information Systems*. In preparing these volumes, leading photogrammetrists from government agencies as well as private and commercial firms and educational institutions have authored and coauthored chapters in their various special areas of expertise. The American Society for Photogrammetry and Remote Sensing also publishes *Photogrammetric Engineering and Remote Sensing*,[3] a monthly journal which brings new developments and applications to the attention of its readers. The society regularly sponsors technical meetings, at various locations throughout the United States, which bring together large numbers of photogrammetrists for the presentation of papers, discussion of new ideas and problems, and firsthand viewing of the latest photogrammetric equipment and software.

The fields of photogrammetry and surveying are so closely knit that it is difficult to separate them. Both are measurement sciences dealing with the production of maps. The American Congress on Surveying and Mapping (ACSM), although primarily concerned with more traditional types of ground surveys, is also vitally interested in photogrammetry. The quarterly journal of ACSM, *Surveying and Land Information Science*, frequently carries photogrammetry-related articles.

The Geomatics Division of the American Society of Civil Engineers (ASCE) is also dedicated to surveying and photogrammetry. Articles on photogrammetry are frequently published in its *Journal of Surveying Engineering*.

The Canadian Institute of Geomatics (CIG) is the foremost professional organization of Canada concerned with photogrammetry. The CIG regularly sponsors technical meetings, and its journal, *Geomatica*, carries photogrammetry articles. The *Journal of Spatial Science* and *Photogrammetric*

Record are similar journals with wide circulation, published in English, by professional organizations in Australia and Great Britain, respectively.

The International Society for Photogrammetry and Remote Sensing (ISPRS), founded in 1910, fosters the exchange of ideas and information among photogrammetrists all over the world. Approximately a hundred foreign countries having professional organizations similar to the American Society for Photogrammetry and Remote Sensing form the membership of ISPRS. This society fosters research, promotes education, and sponsors international conferences at four-year intervals. Its organization consists of seven technical commissions, each concerned with a specialized area in photogrammetry and remote sensing. Each commission holds periodic symposia where photogrammetrists gather to hear presented papers on subjects of international interest. The society's official journal is the *ISPRS Journal of Photogrammetry and Remote Sensing*, which is published in English.

References

American Society for Photogrammetry and Remote Sensing: *Manual of Remote Sensing*, 3d ed., Bethesda, MD, 1998.

———: *Manual of Photographic Interpretation*, 2d ed., Bethesda, MD, 1997.

———: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

———: *Manual of Geographic Information Systems*, Bethesda, MD, 2009.

American Society of Photogrammetry: *Manual of Photogrammetry*, 4th ed., Bethesda, MD, 1980.

Doyle, F. J.: "Photogrammetry: The Next Two Hundred Years," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 5, 1977, p. 575.

Gruner, H.: "Photogrammetry 1776–1976," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 5, 1977, p. 569.

Gutelius, B.: "Engineering Applications of Airborne Scanning Lasers: Reports from the Field," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 4, 1998, p. 246.

Konecny, G.: "Paradigm Changes in ISPRS from the First to the Eighteenth Congress in Vienna," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 10, 1996, p. 1117.

Kraus, K.: *Photogrammetry, Geometry from Images and Laser Scans*, 2d ed., de Gruyter, Berlin, Germany, 2007.

Lillesand, T. M., R. W. Kiefer, and J. W. Chipman: *Remote Sensing and Image Interpretation*, 6th ed., Wiley, New York, 2008.

Merchant, J. W., et al: "Special Issue: Geographic Information Systems," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 11, 1996, p. 1243.

Mikhail, E. M., J. S. Bethel, and J. C. McGlone: *Introduction to Modern Photogrammetry*, Wiley, New York, 2001.

Mikhail, E. M.: "Is Photogrammetry Still Relevant?" *Photogrammetric Engineering and Remote Sensing*, vol. 65, no. 7, 1999, p. 740.

Poore, B. S., and M. DeMulder: "Image Data and the National Spatial Data Infrastructure," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 1, 1997, p. 7.

Ridley, H. M., P. M. Atkinson, P. Aplin, J. P. Muller, and I. Dowman: "Evaluating the Potential of the Forthcoming Commercial U.S. High-Resolution Satellite Sensor Imagery at the Ordnance Survey," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 8, 1997, p. 997.

Rosenblum, N.: *World History of Photography*, Cross River Press, New York, 1989.

Terry, N. G., Jr.: "Field Validation of the UTM Gridded Map," *Photogrammetric Engineering and*

Remote Sensing, vol. 63, no. 4, 1997, p. 381.

Problems

1-1. Explain the differences between metric and interpretative photogrammetry.

1-2. Describe the different classifications of aerial photographs.

1-3. What is the primary difference between high and low oblique aerial photographs?

1-4. Define the following photogrammetric terms: end lap, side lap, stereopair, exposure station, and flying height.

1-5. Discuss some of the principal uses of aerial photogrammetry.

1-6. Discuss some of the principal uses of terrestrial photogrammetry.

1-7. Describe how you would go about obtaining existing aerial photographic coverage of an area.

1-8. To what extent is photogrammetry being used in highway planning in your state?

1-9. Discuss the importance of photogrammetry in geographic information systems.

1-10. Visit the following websites, and briefly discuss the information they provide regarding photogrammetry and mapping.

(*a*) http://www.asprs.org/
(*b*) http://www.isprs.org
(*c*) http://www.cig-acsg.ca/
(*d*) http://www.sssi.org.au/
(*e*) http://www.rspsoc.org
(*f*) http://www.nga.mil/
(*g*) http://www.fgdc.gov/
(*h*) http://www.usgs.gov/pubprod/

---

[1] Information about aerial photographic coverage can be obtained at http://www.usgs.gov/pubprod/aerial.html.

[2] Information about aerial photographic coverage can be obtained at http://www.fsa.usda.gov/fsa.

[3] The title of this journal was changed from *Photogrammetric Engineering* to *Photogrammetric Engineering and Remote Sensing* in 1975.

CHAPTER 2

Principles of Photography and Imaging

2-1 Introduction

Photography, which means "drawing with light," originated long before cameras and light-sensitive photographic films came into use. Ancient Arabs discovered that when inside a dark tent, they could observe inverted images of illuminated outside objects. The images were formed by light rays which passed through tiny holes in the tent. The principle involved was actually that of the pinhole camera of the type shown in Fig. 2-1. In the 1700s French artists used the pinhole principle as an aid in drawing perspective views of illuminated objects. While inside a dark box, they traced the outlines of objects projected onto the wall opposite a pinhole. In 1839 Louis Daguerre of France developed a photographic film which could capture a permanent record of images that illuminated it. By placing this film inside a dark "pinhole box," a picture or *photograph* could be obtained without the help of an artist. This box used in conjunction with photographic film became known as a *camera*. Tremendous improvements have been made in photographic films and film cameras over the years; however, their basic principle has remained essentially unchanged.

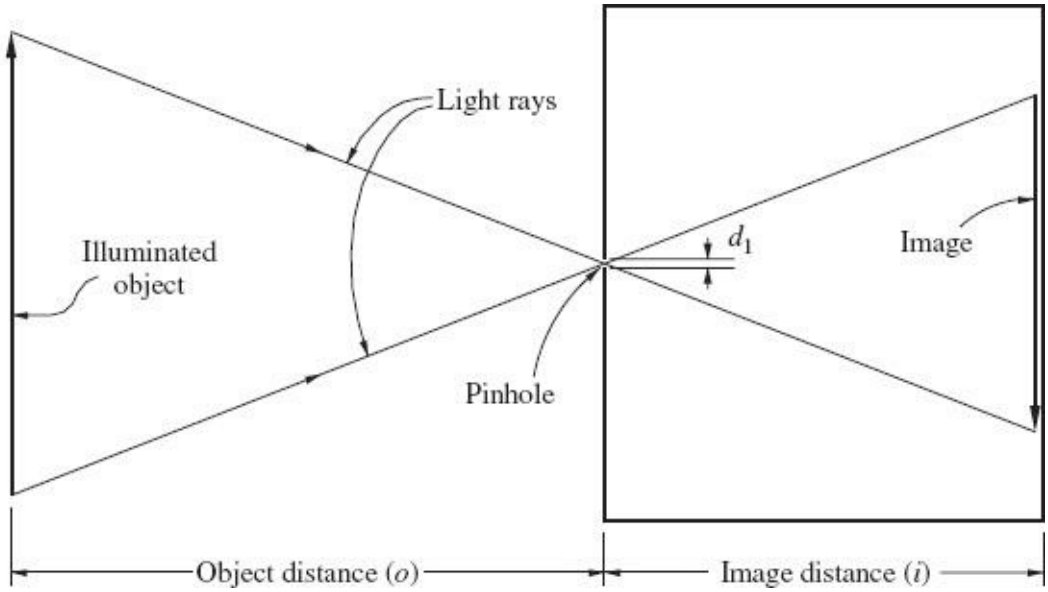


FIGURE 2-1 Principle of the pinhole camera.

A more recent innovation in imaging technology is the digital camera which relies on electronic sensing devices rather than conventional film. The resulting image, called a *digital image*, is stored in computer memory which enables direct computer manipulation.

This chapter describes basic principles that are fundamental to understanding imaging processes. Concepts relevant to both photography and digital imaging are included.

2-2 Fundamental Optics

Both film and digital cameras depend upon optical elements, especially lenses, to function. Thus an

understanding of some of the fundamental principles of optics is essential to the study of photography and imaging.

The science of optics consists of two principal branches: *physical optics* and *geometric optics*. In physical optics, light is considered to travel through a transmitting medium such as air in a series of electromagnetic waves emanating from a point source. Conceptually this can be visualized as a group of concentric circles expanding or *radiating* away from a light source, as illustrated in Fig. 2-2. In nature, a good resemblance of this manner in which light waves propagate can be created by dropping a small pebble into a pool of still water, to create waves radiating from the point where the pebble was dropped. As with water, each light wave has its own *frequency*, *amplitude*, and *wavelength*. Frequency is the number of waves that pass a given point in a unit of time; amplitude is the measure of the height of the crest or depth of the trough; and wavelength is the distance between any wave and the next succeeding one. The speed with which a wave moves from a light source is called its *velocity*. Velocity is related to frequency and wavelength according to the equation

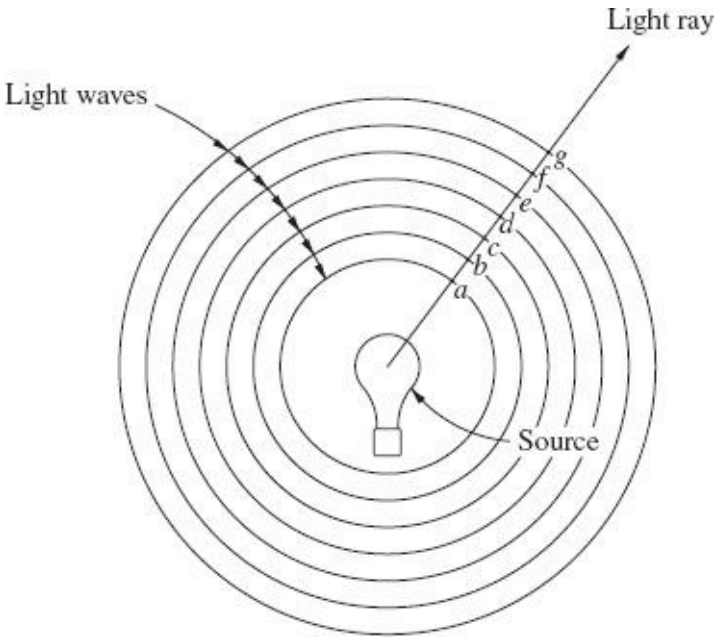


FIGURE 2-2 Light waves emanating from a point source in accordance with the concept of physical optics.

$$V = f\lambda \tag{2-1}$$

In Eq. (2-1), V is velocity, usually expressed in units of meters per second; f is frequency, generally given in cycles per second, or *hertz*; and λ is wavelength, usually expressed in meters. Light has an extremely high velocity, moving at the rate of 2.99792458×10^8 meters per second (m/s) in a vacuum.

In geometric optics, light is considered to travel from a point source through a transmitting medium in straight lines called *light rays*. As illustrated in Fig. 2-3, an infinite number of light rays radiate in all directions from any point source. The entire group of radiating lines is called a *bundle of rays*. This concept of radiating light rays develops logically from physical optics if one considers the travel path of any specific point on a light wave as it radiates away from the source. In Fig. 2-2, for example, point *a* radiates to *b*, *c*, *d*, *e*, *f*, etc. as it travels from the source, thus creating a light ray.

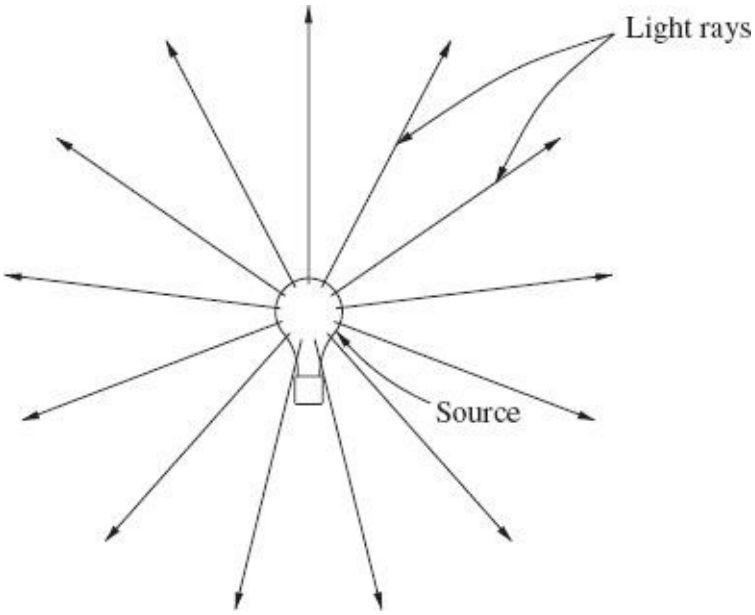


FIGURE 2-3 Bundle of rays emanating from a point source in accordance with the concept of geometric optics.

In analyzing and solving photogrammetric problems, rudimentary line diagrams are often necessary. Their preparation generally requires tracing the paths of light rays through air and various optical elements. These same kinds of diagrams are often used as a basis for deriving fundamental photogrammetric equations. For these reasons, a basic knowledge of the behavior of light, and especially of geometric optics, is prerequisite to a thorough understanding of the science of photogrammetry.

When light passes from one transmitting material to another, it undergoes a change in velocity in accordance with the composition of the substances through which it travels. Light achieves its maximum velocity traveling through a vacuum, it moves more slowly through air, and travels still more slowly through water and glass.

The rate at which light travels through any substance is represented by the *refractive index* of the material. Refractive index is simply the ratio of the speed of light in a vacuum to its speed through a substance, or

$$n = \frac{c}{V} \tag{2-2}$$

In Eq. (2-2), n is the refractive index of a material, c is the velocity of light in a vacuum, and V is its velocity in the substance. The refractive index for any material, which depends upon the wavelength of the light, is determined through experimental measurement. Typical values for indexes of refraction of common media are vacuum, 1.0000; air, 1.0003; water, 1.33; and glass, 1.5 to 2.0.

When light rays pass from one homogeneous, transparent medium to a second such medium having a different refractive index, the path of the light ray is bent or *refracted*, unless it intersects the second medium normal to the interface. If the intersection occurs obliquely, as shown in Fig. 2-4, then the *angle of incidence*, ϕ, is related to the *angle of refraction*, ϕ', by the law of refraction, frequently called *Snell's law*. This law is stated as follows:

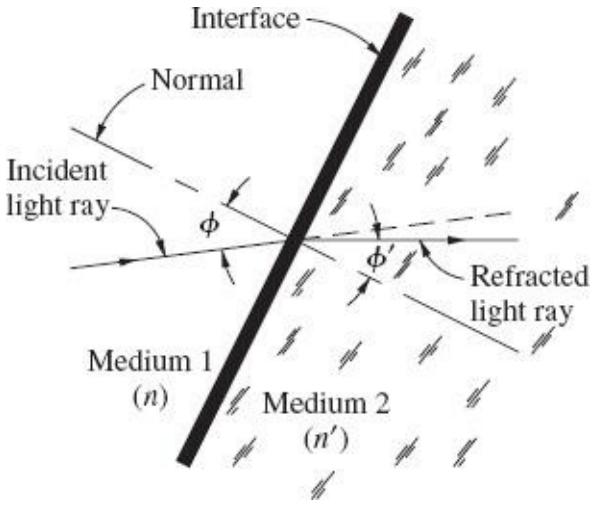


FIGURE 2-4 Refraction of light rays.

$$n \sin \phi = n' \sin \phi' \tag{2-3}$$

where n is the refractive index of the first medium and n' is the refractive index of the second medium. The angles ϕ and ϕ' are measured from the normal to the incident and refracted rays, respectively.

Light rays can also be made to change directions by reflection. When a light ray strikes a smooth surface such as a highly polished metal mirror, it is reflected so that the *angle of reflection* ϕ'' is equal to the incidence angle ϕ, as shown in Fig. 2-5*a*. Both angles lie in a common plane and are measured from NN', the normal to the reflecting surface.

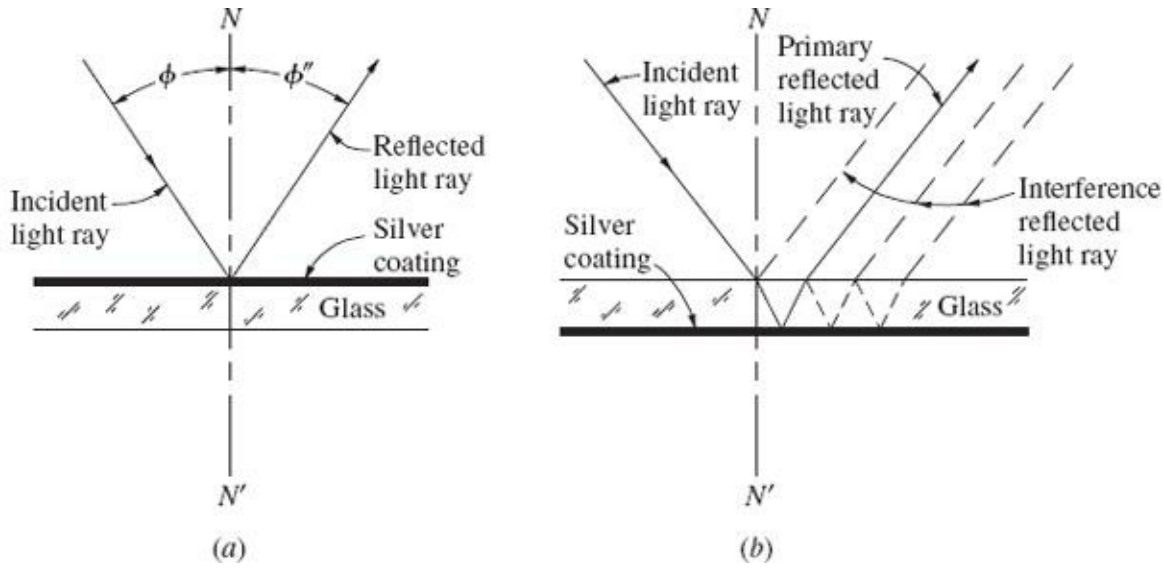


FIGURE 2-5 (*a*) First-surface mirror demonstrating the angle of incidence ϕ and angle of refection ϕ''; (*b*) back-surfaced mirror.

Plane mirrors used for nonscientific purposes generally consist of a plane sheet of glass with a thin reflective coating of silver on the back. This type of "back-surfaced" mirror is optically undesirable, however, because it creates multiple reflections that interfere with the primary reflected light ray, as shown in Fig. 2-5*b*. These undesirable reflections may be avoided by using *first-surface* mirrors, which have their silver coating on the front of the glass, as shown in Fig. 2-5*a*.

2-3 Lenses

A simple lens consists of a piece of optical glass that has been ground so that it has either two spherical surfaces or one spherical surface and one flat surface. Its primary function is to gather light rays from object points and bring them to focus at some distance on the opposite side of the lens. A lens accomplishes this function through the principles of refraction. The simplest and most primitive device that performs the functions of a lens is a tiny pinhole which theoretically allows a single light ray from each object point to pass. The tiny hole of diameter d_1 of the pinhole camera illustrated in Fig. 2-1 produces an inverted image of the object. The image is theoretically in focus regardless of the distance from the pinhole to the camera's image plane. Pinholes allow so little light to pass, however, that they are unsuitable for photogrammetric work. For practical purposes they are replaced with larger openings occupied by glass lenses.

The advantage of a lens over a pinhole is the increased amount of light that is allowed to pass. A lens gathers an entire *pencil of rays* from each object point instead of only a single ray. As discussed earlier and illustrated in Fig. 2-3, when an object is illuminated, each point in the object reflects a bundle of light rays. This condition is also illustrated in Fig. 2-6. A lens placed in front of the object gathers a pencil of light rays from each point's bundle of rays and brings these rays to focus at a point in a plane on the other side of the lens, called the *image plane*. An infinite number of image points, focused in the image plane, form the image of the entire object. Note from Fig. 2-6 that the image is inverted by the lens.

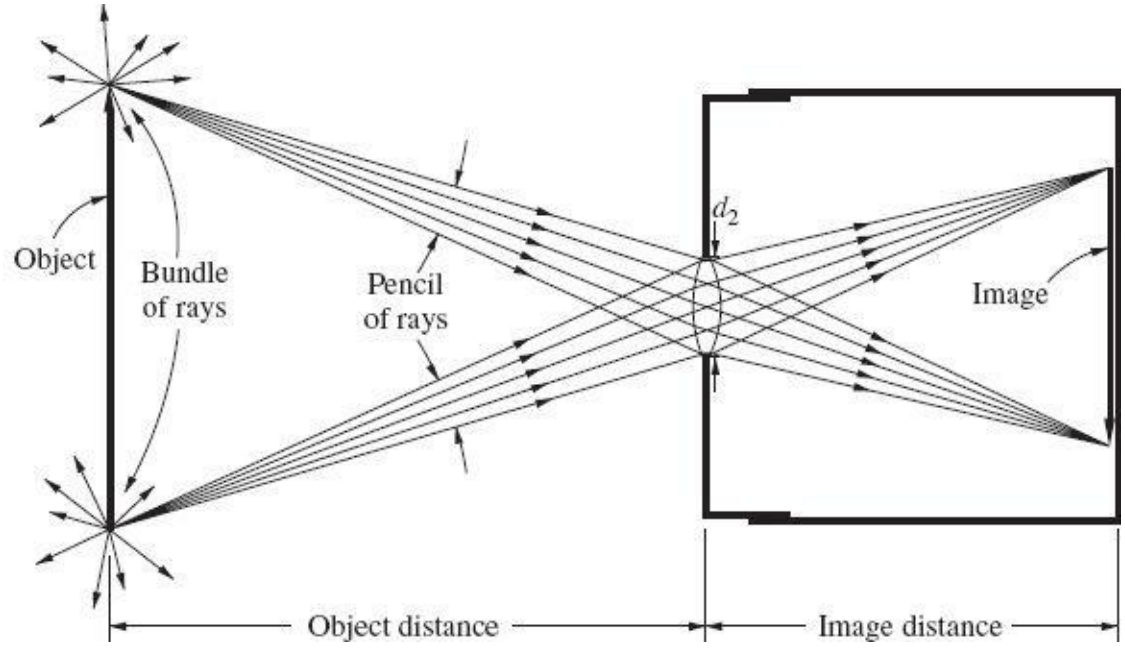


FIGURE 2-6 Pencils of light and image formation in a single-lens camera.

The *optical axis* of a lens is defined as the line joining the centers of curvature of the spherical surfaces of the lens (points O_1 and O_2 of Fig. 2-7). In this figure R_1 and R_2 are the radii of the lens surfaces, and the optical axis is the line O_1O_2. Light rays that are parallel to the optical axis as they enter a lens come to focus at F, the *focal point* of the lens. The distance from the focal point to the center of a lens is f, the *focal length* of the lens. A plane perpendicular to the optical axis passing through the focal point is called the *plane of infinite focus*, or simply the *focal plane*. Parallel rays entering a *converging lens* (one with two convex exterior surfaces, as shown in Fig. 2-7), regardless of the angle they make with the optical axis, are ideally brought to focus in the plane of infinite focus

(see the dashed rays of the figure).

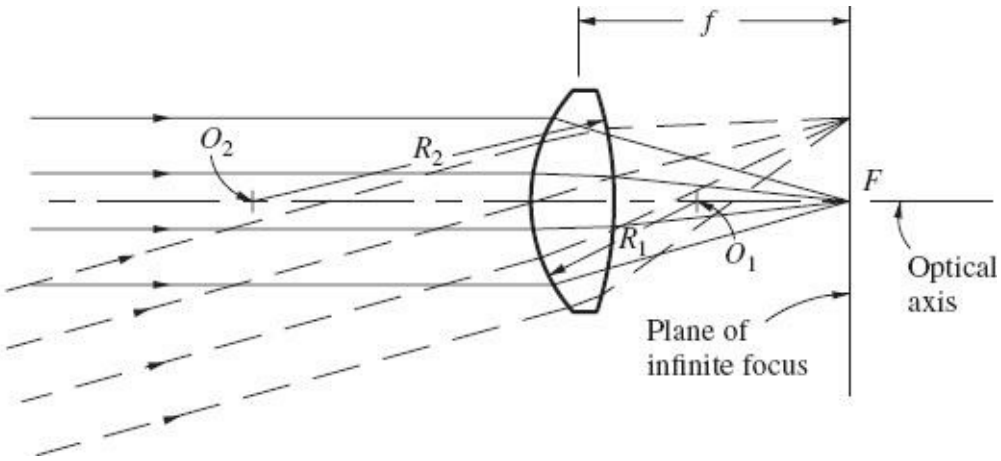


FIGURE 2-7 Optical axis, focal length, and plane of infinite focus of a lens.

Example 2-1

A single ray of light traveling through air (n = 1.0003) enters a convex glass lens (n' = 1.52) having a radius of 5.00 centimeters (cm), as shown in Fig. 2-8. If the light ray is parallel to and 1.00 cm above the optical axis of the lens, what are the angles of incidence ϕ and refraction ϕ' for the air-to-glass interface?

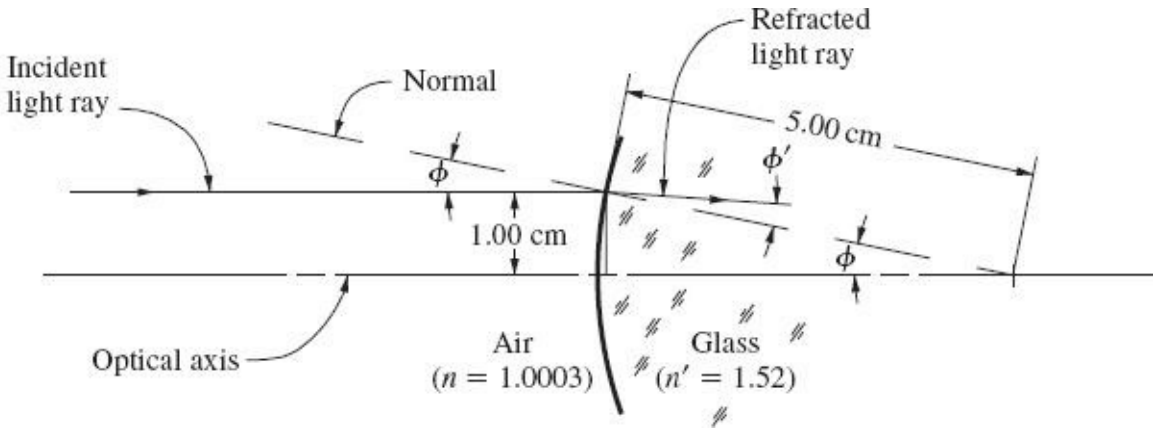


FIGURE 2-8 Refraction of an incident light ray parallel to the optical axis of a lens.

Solution From the figure,

$$\sin\phi = \frac{1.00\text{ cm}}{5.00\text{ cm}} \qquad \text{from which} \qquad \phi = 11.5°$$

Applying Snell's law [Eq. (2-3)] gives

$$n \sin\phi = n' \sin\phi'$$

$$1.0003\left(\frac{1.00\text{ cm}}{5.00\text{ cm}}\right) = 1.52 \sin\phi' \qquad \text{from which} \qquad \phi' = 7.56° \quad ▲$$

A pencil of incident light rays coming from an object located an infinite distance away from the lens will be parallel, as illustrated in Fig. 2-7, and the image will come to focus in the plane of infinite focus. For objects located some finite distance from the lens, the *image distance* (distance from lens center to image plane) is greater than the focal length. The following equation, called the *lens formula*, expresses the relationship of object distance o and image distance i to the focal length f of a converging lens:

$$\frac{1}{o}+\frac{1}{i}=\frac{1}{f} \tag{2-4}$$

If the focal length of a lens and the distance to an object are known, the resulting distance to the image plane can be calculated by using the lens formula.

Example 2-2
Find the image distance for an object distance of 50.0 m and a focal length of 50.0 cm.

Solution By Eq. (2-4),

$$\frac{1}{50.0}+\frac{1}{i}=\frac{1}{0.500}$$

$$\frac{1}{i}=\frac{1}{0.500}-\frac{1}{50.0}=2.00-0.0200=1.98$$

$$i=\frac{1}{1.98}=0.505\text{ m}=50.5\text{ cm}\quad\blacktriangle$$

The preceding analysis of lenses was simplified by assuming that their thicknesses were negligible. With *thick lenses*, this assumption is no longer valid. Thick lenses may consist of a single thick element or a combination of two or more elements which are either cemented together in contact or otherwise rigidly held in place with airspaces between the elements. A thick "combination" lens used in an aerial camera is illustrated in Fig. 2-9. Note that it consists of 15 individual elements.

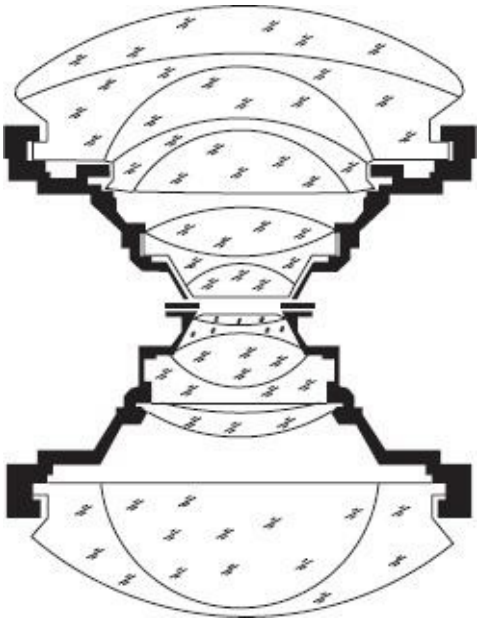


FIGURE 2-9 Cross section of SAGA-F lens. (*Drawing from brochure courtesy of LH Systems, LLC.*)

Two points called *nodal points* must be defined for thick lenses. These points, termed the *incident* nodal point and the *emergent* nodal point, lie on the optical axis. They have the property that conceptually, any light ray directed toward the incident nodal point passes through the lens and emerges on the other side in a direction parallel to the original incident ray and directly away from the emergent nodal point. In Fig. 2-10, for example, rays *AN* and *N'a* are parallel, as are rays *BN* and *N'b*. Points *N* and *N'* are the incident and emergent nodal points, respectively, of the thick lens. Such light rays do not necessarily pass through the nodal points, as illustrated by the figure.

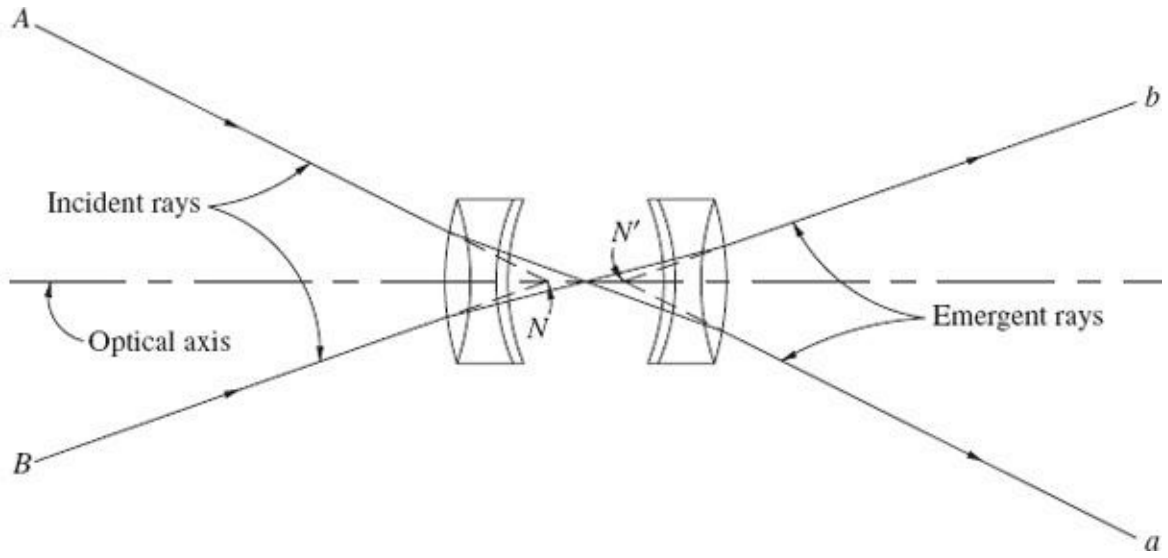


FIGURE 2-10 Nodal points of a thick lens.

If parallel incident light rays (rays from an object at an infinite distance) pass through a thick lens, they will come to focus at the plane of infinite focus. The focal length of a thick lens is the distance from the emergent nodal point *N'* to this plane of infinite focus.

It is impossible for a single lens to produce a perfect image; it will, instead, always be somewhat blurred and geometrically distorted. The imperfections that cause blurring, or degrade the sharpness of the image, are termed *aberrations*. Through the use of additional lens elements, lens designers are able to correct for aberrations and bring them within tolerable limits. *Lens distortions*, on the other hand, do not degrade image quality but deteriorate the geometric quality (or positional accuracy) of the image. Lens distortions are classified as either *symmetric radial* or *decentering*. Both occur if light rays are bent, or change directions, so that after they pass through the lens, they do not emerge parallel to their incoming directions. Symmetric radial distortion, as its name implies, causes imaged points to be distorted along radial lines from the optical axis. Outward radial distortion is considered positive, and inward radial distortion is considered negative. Decentering distortion, which has both tangential and asymmetric radial components, causes an off-center distortion pattern. These lens distortions are discussed in greater detail in Secs. 3-10 and 4-11.

Resolution or *resolving power* of a lens is the ability of the lens to show detail. One common method of measuring lens resolution is to count the number of *line pairs* (black lines separated by white spaces of equal thickness) that can be clearly distinguished within a width of 1 millimeter (mm) in an image produced by the lens. The *modulation transfer function* (MTF) is another way of specifying the resolution characteristics of a lens. Both methods are discussed in Sec. 3-14, and a line-

pair test pattern is shown in Fig. 3-19. Good resolution is important in photogrammetry because photo images must be sharp and clearly defined for precise measurements and accurate interpretative work. Photographic resolution is not just a function of the camera lens, however, but also depends on other factors, as described in later sections of this book.

The *depth of field* of a lens is the range in object distance that can be accommodated by a lens without introducing significant image deterioration. For a given lens, depth of field can be increased by reducing the size of the lens opening (*aperture*). This limits the usable area of the lens to the central portion. For aerial photography, depth of field is seldom of consequence, because variations in the object distance are generally a very small percentage of the total object distance. For close-range photography, however (see Chap. 19), depth of field is often extremely critical. The shorter the focal length of a lens, the greater its depth of field, and vice versa. Thus, if depth of field is critical, it can be somewhat accommodated either through the selection of an appropriate lens or by reducing aperture size.

Vignetting and *falloff* are lens characteristics which cause resultant images to appear brighter in the center than around the edges. Compensation can be provided for these effects in the lens design itself, by use of an antivignetting filter in the camera, or through lighting adjustments in the printing process (see Sec. 2-8).

2-4 Single-Lens Camera

One of the most fundamental instruments used in photogrammetry is the single-lens camera. The geometry of this device, depicted in Fig. 2-6, is similar to that of the pinhole camera, shown in Fig. 2-1. In the single-lens camera, the size of the aperture (d_2 in Fig. 2-6) is much larger than a pinhole, requiring a lens in order to maintain focus.

Instead of object distances and image distances being unrestricted as with the pinhole camera, with the lens camera these distances are governed by the lens formula, Eq. (2-4). To satisfy this equation, the lens camera must be focused for each different object distance by adjusting the image distance. When object distances approach infinity, such as for photographing objects at great distances, the term $1/o$ in Eq. (2-4) approaches *zero* and image distance i is then equal to f, the lens focal length. With aerial photography, object distances are very great with respect to image distances; therefore aerial cameras are manufactured with their focus fixed for infinity. This is accomplished by fixing image distance equal to the focal length of the camera lens.

2-5 Illuminance

Illuminance of any photographic exposure is the brightness or amount of light received per unit area on the image plane surface during exposure. A common unit of illuminance is the *meter-candle*. One meter-candle (1 m · cd) is the illuminance produced by a standard candle at a distance of 1 m.

Illuminance is proportional to the amount of light passing through the lens opening during exposure, and this is proportional to the area of the opening. Since the area of the lens opening is $\pi d^2/4$, illuminance is proportional to the variable d^2, the square of the diameter of the lens opening.

Image distance i is another factor which affects illuminance. Illuminance is an effect that adheres to the *inverse square law*, which means that the amount of illuminance is inversely proportional to the square of distance from the aperture. According to this law, at the center of the photograph, illuminance is proportional to $1/i^2$. As distances increase away from the center of the photograph, distances from the aperture likewise increase. This causes decreased illuminance, an effect which can

be quite severe for wide-angle lenses. This is one aspect of the physical basis for lens falloff, mentioned in Sec. 2-3. Normally in photography, object distances are sufficiently long that the term $1/o$ in Eq. (2-4) is nearly zero, in which case i is equal to f. Thus, at the center of a photograph, illuminance is proportional to the quantity $1/f^2$, and the two quantities may be combined so that illuminance is proportional to d^2/f^2. The square root of this term is called the *brightness factor*, or

$$\sqrt{\frac{d^2}{f^2}} = \frac{d}{f} = \text{brightness factor} \tag{2-5}$$

The inverse of Eq. (2-5) is also an inverse expression of illuminance and is the very common term *f-stop*, also called *f-number*. In equation form, According to Eq. (2-6), *f*-stop is the ratio of focal length to the diameter of the lens opening, or *aperture*. As the aperture increases, *f*-stop numbers decrease and illuminance increases, thus requiring less exposure time, i.e., faster shutter speeds. Because of this correlation between *f*-stop and shutter speed, *f*-stop is the term used for expressing *lens speed* or the "light-gathering" power of a lens. Illuminance produced by a particular lens is correctly expressed by Eq. (2-6), whether the lens has a very small diameter with short focal length or a very large diameter with a long focal length. If *f*-stop is the same for two different lenses, the illuminance at the center of each of their images will be the same.

$$f\text{-stop} = \frac{f}{d} \tag{2-6}$$

2-6 Relationship of Aperture and Shutter Speed

A sunbather gets a suntan or sunburn when exposed to sunshine. The darkness of tan or severity of burn is a function of the sun's brightness and the time of exposure to the sun. *Total exposure* of photographic film is likewise the product of illuminance and time of exposure. Its unit is *meter-candle-seconds*, although in certain applications a different unit, the *microlux-second* is used.

In making photographic exposures, the correct amounts of illuminance and time may be correlated using a *light meter*. Illuminance is regulated by varying *f*-stop settings on the camera, while time of exposure is set by varying the shutter speed. Variations in *f*-stop settings are actually variations in the diameter of the aperture, which can be controlled with a *diaphragm*—a circular shield that enlarges or contracts, changing the diameter of the opening of the lens and thus regulating the amount of light that is allowed to pass through the lens.

With a lens camera, as the diameter of the aperture increases, enabling faster exposures, the depth of field becomes less and lens distortions become more severe. At times a small diaphragm opening is desirable, and there are times when the reverse is true. To photograph a scene with great variations in object distances and yet retain sharp focus of all images, a large depth of field is required. In this case, to maximize depth of field, the picture would be taken at a slow shutter speed and large *f*-stop setting, corresponding to a small-diameter lens opening. On the other hand, in photographing rapidly moving objects or in making exposures from a moving vehicle such as an airplane, a fast shutter speed is essential, to reduce image motion. In this situation a small *f*-stop setting corresponding to a large-diameter lens opening would be necessary for sufficient exposure.

From the previous discussion it is apparent that there is an important relationship between *f*-stop and shutter speed. If exposure time is cut in half, total exposure is also halved. Conversely, if aperture

area is doubled, total exposure is doubled. If shutter time is halved and aperture area is doubled, total exposure remains unchanged.

Except for inexpensive models, cameras are manufactured with the capability of varying both shutter speed and *f*-stop setting, and many modern cameras do this function automatically. The nominal *f*-stop settings are 1, 1.4, 2.0, 2.8, 4.0, 5.6, 8.0, 11, 16, 22, and 32. Not all cameras have all these, but the more expensive cameras have many of them. The camera pictured in Fig. 2-11, for example, has a minimum *f*-stop setting of *f*-2.8 and is also equipped for varying shutter speeds down to $\frac{1}{4000}$ second (s).

FIGURE 2-11 Digital single-lens reflex camera having a minimum *f*-stop setting of *f*-2.8 and variable shutter speeds ranging down to $\frac{1}{4000}$ s. (*Courtesy University of Florida*)

An *f*-stop number 1, or *f*-1, occurs, according to Eq. (2-6), when the aperture diameter equals the lens focal length. A setting at *f*-1.4 halves the aperture area from that of *f*-1. In fact, each succeeding number of the nominal *f*-stops listed previously halves the aperture area of the preceding one, and it is seen that each succeeding number is obtained by multiplying the preceding one by $\sqrt{2}$. This is illustrated as follows:

Let $d_1 = f$, where d_1 is the aperture diameter. Then

$$\frac{f}{d_1} = 1 = f\text{-stop}$$

At *f*-stop = 1,

$$\text{Aperture area} = A_1 = \frac{\pi(d_1)^2}{4}$$

If the aperture diameter is reduced to d_2, giving a lens opening area of one-half of A_1, then

$$A_2 = \frac{A_1}{2} = \frac{\pi(d_2)^2}{4} = \frac{\pi(d_1)^2}{2(4)}$$

From the above, $d_2 = d_1/\sqrt{2}$ and the corresponding *f*-stop number is

$$f\text{-stop} = \frac{f\sqrt{2}}{d_1} = 1\sqrt{2} = 1.4$$

The relationship between *f*-stop and shutter speed leads to many interesting variations in obtaining correct exposures. Many digital cameras have automatic controls that will set the *f*-stop and shutter speed for proper exposure. In addition to a manual mode, they typically provide: (1) a fully automatic mode, where both *f*-stop and shutter speed are appropriately selected, (2) an aperture priority mode, where the user inputs a fixed *f*-stop and the camera selects the appropriate shutter speed, and (3) a shutter priority mode, where the user inputs a fixed shutter speed and the camera selects the appropriate *f*-stop.

Example 2-3

Suppose that a photograph is optimally exposed with an *f*-stop setting of *f*-4 and a shutter speed of $\frac{1}{500}$ s. What is the correct *f*-stop setting if shutter speed is changed to $\frac{1}{1000}$ s?

Solution Total exposure is the product of diaphragm area and shutter speed. This product must remain the same for the $\frac{1}{1000}$ -s shutter speed as it was for the $\frac{1}{500}$ -s shutter speed, or

$$\text{Area}_1 \times \text{time}_1 = \text{area}_2 \times \text{time}_2$$

Rearranging, we have

$$\text{Area}_2 = \text{area}_1 \times \frac{\text{time}_1}{\text{time}_2}$$

(*a*)

Let d_1 and d_2 be diaphragm diameters for $\frac{1}{500}$ - and $\frac{1}{1000}$ -s shutter times, respectively. Then the respective diaphragm areas are

$$\text{Area}_1 = \frac{\pi(d_1)^2}{4} \qquad \text{and} \qquad \text{area}_2 = \frac{\pi(d_2)^2}{4}$$

(*b*)

By Eq. (2-6),

$$d_1 = \frac{f}{f\text{-stop}_1} \qquad \text{and since } f\text{-stop}_1 = 4 \qquad d_1 = \frac{f}{4}$$

(*c*)

Substituting (*b*) and (*c*) into (*a*) gives

$$\frac{\pi(d_2)^2}{4} = \frac{\pi(f)^2}{4(4)^2} \times \frac{1/500}{1/1000}$$

Reducing gives

$$\frac{f}{d_2} = \sqrt{\frac{500 \times 16}{1000}} = 2.8$$

Hence *f*-2.8 is the required *f*-stop. The above is simply computational proof of an earlier statement that each successive nominal *f*-stop setting halves the aperture area of the previous one; or in this case *f*-2.8 doubles the aperture area of *f*-4, which is necessary to retain the same exposure if shutter time is halved. ▲

2-7 Characteristics of Photographic Emulsions

Photographic films consist of two parts: *emulsion* and *backing* or *support*. The emulsion contains light-sensitive silver halide crystals. These are placed on the backing or support in a thin coat, as shown in Fig. 2-12. The support material is usually paper, plastic film, or glass.

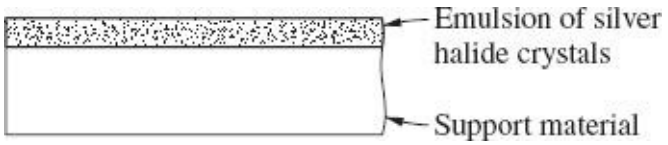


FIGURE 2-12 Cross section of a photographic film.

When silver halide crystals are exposed to light, the bond between the silver and the halide is

weakened. An emulsion that has been exposed to light contains an invisible image of the object, called the *latent image*. When the latent image is developed, areas of the emulsion that were exposed to intense light turn to free silver and become black. Areas that received no light become white if the support is white paper. (They become clear if the support is glass or transparent plastic film.) The degree of darkness of developed images is a function of the total exposure (product of illuminance and time) that originally sensitized the emulsion to form the latent image. In any photographic exposure, there will be variations in illuminance received from different objects in the photographed scene, and therefore between black and white there will exist various tones of gray which result from these variations in illuminance. Actually the crystals turn black, not gray, when exposed to sufficient light. However, if the light received in a particular area is sufficient to sensitize only a portion of the crystals, then a gray tone results from a mixture of the resulting black and white. The greater the exposure, the greater the percentage of black in the mixture and hence the darker the shade of gray.

The degree of darkness of a developed emulsion is called its *density*. The greater the density, the darker the emulsion. Density of a developed emulsion on a transparent film can be determined by subjecting the film to a light source, and then comparing the intensity of incident light upon the film to that which passes through (transmitted light). The relationship is expressed in Eq. (2-7), where D is the density. Since the intensity response of a human eye is nonlinear, the base-ten logarithm (log) is used so that density will be nearly proportional to perceived brightness. A density value of zero corresponds to a completely transparent film, whereas a film that allows 1 percent of the incident light to pass through has a density of 2. The amount of light incident to an emulsion and the amount transmitted can be measured with an instrument called a *densitometer*.

$$D = \log\left(\frac{\text{incident intensity}}{\text{transmitted intensity}}\right) \tag{2-7}$$

If exposure is varied for a particular emulsion, corresponding variations in densities will be obtained. A plot of density on the ordinate versus logarithm of exposure on the abscissa for a given emulsion produces a curve called the *characteristic curve*, also known as the *D–log E curve*, or the *H-and-D curve*. A typical characteristic curve is shown in Fig. 2-13. Characteristic curves for different emulsions vary somewhat, but they all have the same general shape. The lower part of the curve, which is concave upward, is known as the *toe* region. The upper portion, which is concave downward, is the *shoulder* region. A *straight-line* portion occurs between the toe and shoulder regions.

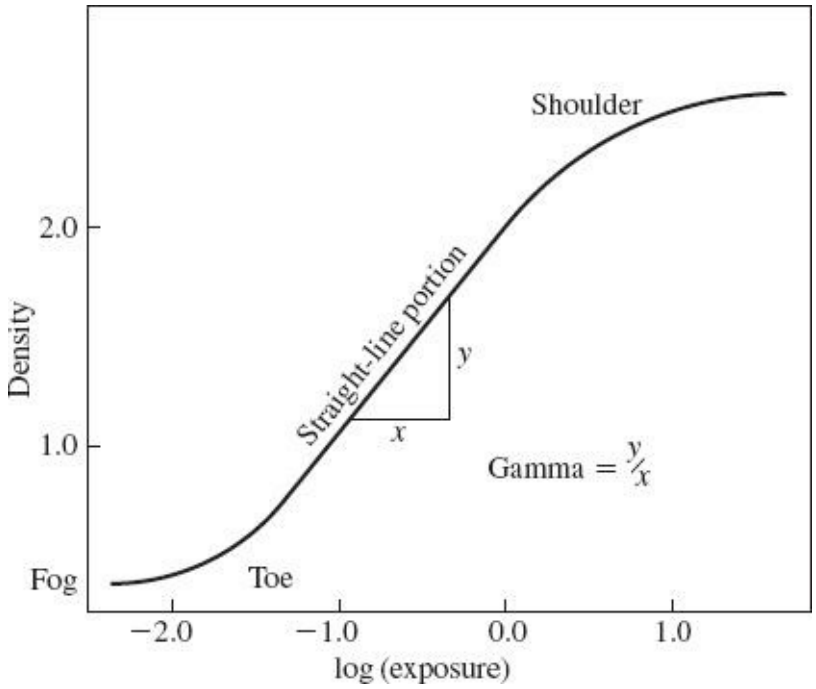


FIGURE 2-13 Typical "characteristic curve" of a photographic emulsion.

Characteristic curves are useful in describing the characteristics of photographic emulsions. The slope of the straight-line portion of the curve for example, is a measure of the contrast of the film. The steeper the slope, the greater the *contrast* (change in density for a given range of exposure). Contrast of a given film is expressed as *gamma*, the slope of the straight-line portion of the curve, as shown on Fig. 2-13. From the figure it is evident that for an exposure of zero the film has some density. The density of an unexposed emulsion is called *fog*, and on the curve it is the density corresponding to the low portion of the toe region. It is also apparent from Fig. 2-13 that exposure must exceed a certain minimum before density greater than fog occurs. Also, exposures within the shoulder region affect the density very little, if any. Thus, a properly exposed photograph is one in which the entire range of exposure occurs within the straight-line portion of the curve.

Just as the skin of different people varies in sensitivity to sunlight, so does the sensitivity of emulsions vary. Light sensitivity of photographic emulsions is a function of the size and number of silver halide crystals or *grains* in the emulsion. When the required amount of light exposes a grain in the emulsion, the entire grain becomes exposed regardless of its size. If a certain emulsion is composed of grains smaller than those in another emulsion, such that approximately twice as many grains are required to cover the film, then this emulsion will also require about twice as much light to properly expose *the emulsion*. Conversely, as grain size increases, the total number of grains in the emulsion decreases and the amount of light required to properly expose the emulsion decreases. Film is said to be more sensitive and faster when it requires less light for proper exposure. Faster films can be used advantageously in photographing rapidly moving objects.

As sensitivity and grain size increase, the resulting image becomes coarse and *resolution* (sharpness or crispness of the picture) is reduced. Thus, for highest pictorial quality, such as portrait work, slow, finegrained emulsions are preferable. Film resolution can be tested by photographing a standard test pattern, as previously discussed in Sec. 2-3 and covered in greater detail in Sec. 3-14.

Photographers have developed exposure guides for films of various sensitivities. For films not used in aerial photography, the *International Standards Organization* (ISO) number is used to indicate film sensitivity or speed. The ISO number assigned to a film is roughly equal to the inverse of shutter speed (in seconds) required for proper exposure in pure sunlight for a lens opening of *f*-16. According

to this rule of thumb, if a film is properly exposed in pure sunlight at *f*-16 and $\frac{1}{200}$ s, it is classified ISO 200. This rule of thumb is seldom needed today because of the availability of microprocessor-controlled cameras which, given the ISO rating of the film being used, automatically yield proper exposures (*f*-stops and shutter speeds) for particular lighting conditions.

The foregoing discussion of film speed applies primarily to ordinary ground-based photography. In aerial photography, the range of illuminance at the focal plane is significantly lower due to the narrower range of ground illuminance and atmospheric haze. For this reason, the sensitivity of films used in aerial photography is expressed as *aerial film speed* (AFS), which is different from, and should not be confused with, the ISO number. Aerial film speed is determined by the point on the characteristic curve where density is 0.3 unit above the fog density.

2-8 Processing and Printing Black-and-White Photographs

The five-step darkroom procedure for processing an exposed black-and-white emulsion is as follows:

1. *Developing*. The exposed emulsion is placed in a chemical solution called *developer*. The action of the developer causes grains of silver halide that were exposed to light to be reduced to free black silver. The free silver produces the blacks and shades of gray of which the image is composed. Developers vary in strength and other characteristics and must therefore be carefully chosen to produce the desired results.
2. *Stop bath*. When proper darkness and contrast of the image have been attained in the developing stage, it is necessary to stop the developing action. This is done with a stop bath—an *acidic* solution which neutralizes the *basic* developer solution.
3. *Fixing*. Not all the silver halide grains are turned to free black silver as a result of developing. Instead, there remain many undeveloped grains which would also turn black upon exposure to light if they were not removed. To prevent further developing which would ruin the image, the undeveloped silver halide grains are dissolved out in the fixing solution.
4. *Washing*. The emulsion is washed in clean running water to remove any remaining chemicals. If not removed, these chemicals could cause spotting or haziness of the image.
5. *Drying*. The emulsion is dried to remove the water from the emulsion and backing material.

Modern equipment is capable of automatically performing the entire five-step darkroom procedure nonstop. The result obtained from developing black-and-white film is a *negative*. It derives its name from the fact that it is reversed in tone and geometry from the original scene that was photographed; i.e., black objects appear white and vice versa, and images are inverted.

A *positive* print is obtained by passing light through the negative onto another emulsion. This reverses tone and geometry again, thereby producing an image in which those two characteristics are true. The configuration involved in this process may be either *contact* or *projection printing*. In contact printing the emulsion side of a negative is placed in direct contact with the unexposed emulsion contained on printing material. Together these are placed in a contact printer and exposed with the emulsion of the positive facing the light source. Figure 2-14 is a schematic of a single-frame contact printer. In contact printing, the positive that is obtained is the same size as the negative from which it was made.

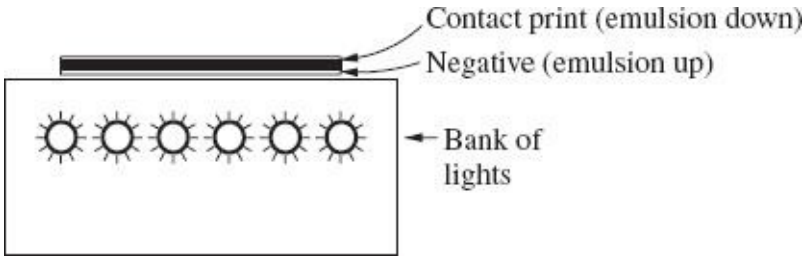


FIGURE 2-14 A contact printer.

A processed negative will generally have a nonuniform darkness due to lens falloff, vignetting, and other effects. Consequently, uniform lighting across the negative during exposure of a positive print will underexpose the emulsion in darker areas of the negative and overexpose it in lighter areas. Compensation for this can be made in a process called *dodging*. It consists of adjusting the amount of light passing through different parts of the negative so that optimum exposure over the entire print is obtained in spite of darkness variations.

If positives are desired at a scale either enlarged or reduced from the original negative size, a *projection printing* process can be used. The geometry of projection printing is illustrated in Fig. 2-15. In this process, the negative is placed in the projector of the printer and illuminated from above. Light rays carry images *c* and *d*, for example, from the negative, through the projector lens, and finally to their locations *C* and *D* on the positive, which is situated on the easel plane beneath the projector. The emulsion of the positive, having been exposed, is then processed in the manner previously described.

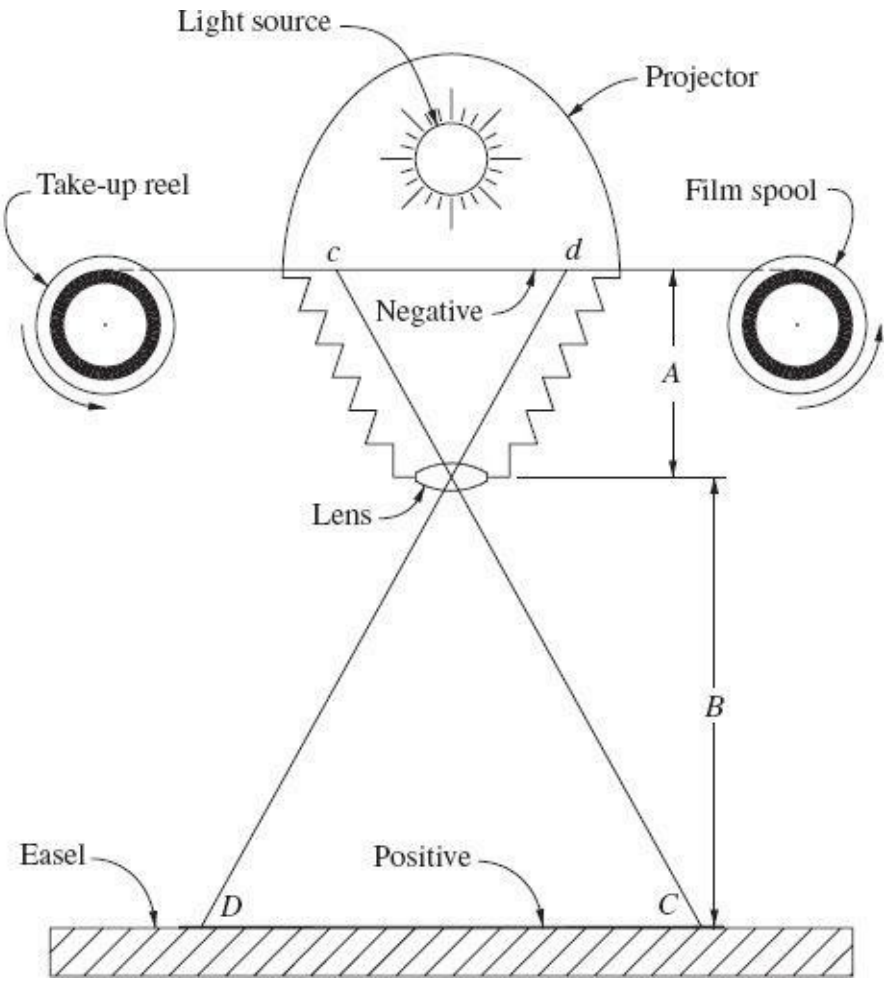


FIGURE 2-15 Geometry of enlargement with a projection printer.

Distances *A* and *B* of Fig. 2-15 can be varied so that positives can be printed at varying scales, and at the same time the lens formula, Eq. (2-4), can be satisfied for the projector's lens. The enlargement or reduction ratio from negative to positive size is equal to the ratio *B*/*A*.

Besides using printing paper, positives may also be prepared on plastic film or glass plates. In photogrammetric terminology, positives prepared on glass plates or transparent plastic materials are called *diapositives*.

2-9 Spectral Sensitivity of Emulsions

The sun and various artificial sources such as lightbulbs emit a wide range of *electromagnetic energy*. The entire range of this electromagnetic energy is called the *electromagnetic spectrum.* X-rays, visible light rays, and radio waves are some familiar examples of energy variations within the electromagnetic spectrum. Electromagnetic energy travels in sinusoidal oscillations called *waves*. Variations in electromagnetic energy are classified according to variations in their wavelengths or frequencies of propagation. The velocity of electromagnetic energy in a vacuum is constant and is related to frequency and wavelength through the following expression (see also Sec. 2-2):

$$c = f\lambda \tag{2-8}$$

In Eq. (2-8), c is the velocity of electromagnetic energy in a vacuum, f is frequency, and λ is wavelength. Figure 2-16 illustrates the wavelength classification of the electromagnetic spectrum. Visible light (that electromagnetic energy to which our eyes are sensitive) is composed of only a very small portion of the spectrum (see the figure). It consists of energy with wavelengths in the range of from about 0.4 to 0.7 micrometer (μm). Energy having wavelengths slightly shorter than 0.4 μm is called *ultraviolet*, and energy with wavelengths slightly longer than 0.7 μm is called *near-infrared*. Ultraviolet and near-infrared cannot be detected by the human eye.

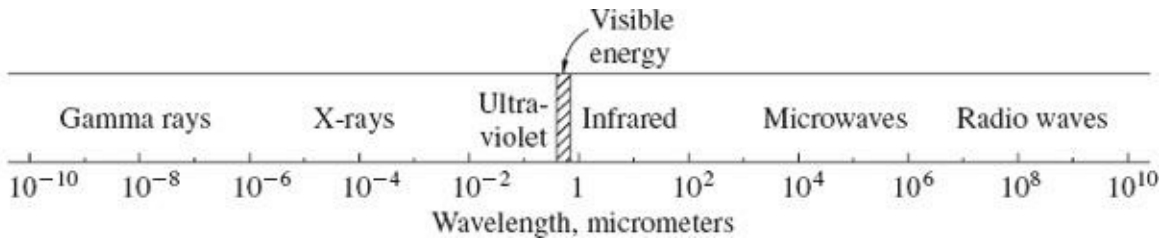


FIGURE 2-16 Classification of the electromagnetic spectrum by wavelength.

Within the wavelengths of visible light, the human eye is able to distinguish different colors. The primary colors—blue, green, and red—are composed of slightly different wavelengths: Blue is composed of energy having wavelengths of about 0.4 to 0.5 μm, green from 0.5 to 0.6 μm, and red from 0.6 to 0.7 μm. To the human eye, other hues can be represented by combinations of the primary colors; e.g., yellow is perceived when red and green light are combined. There are multitudes of these combination colors. White light is the combination of all the visible colors. It can be broken down into its component colors by passing it through a prism, as shown in Fig. 2-17. Color separation occurs because of different refractions that occur with energy of different wavelengths.

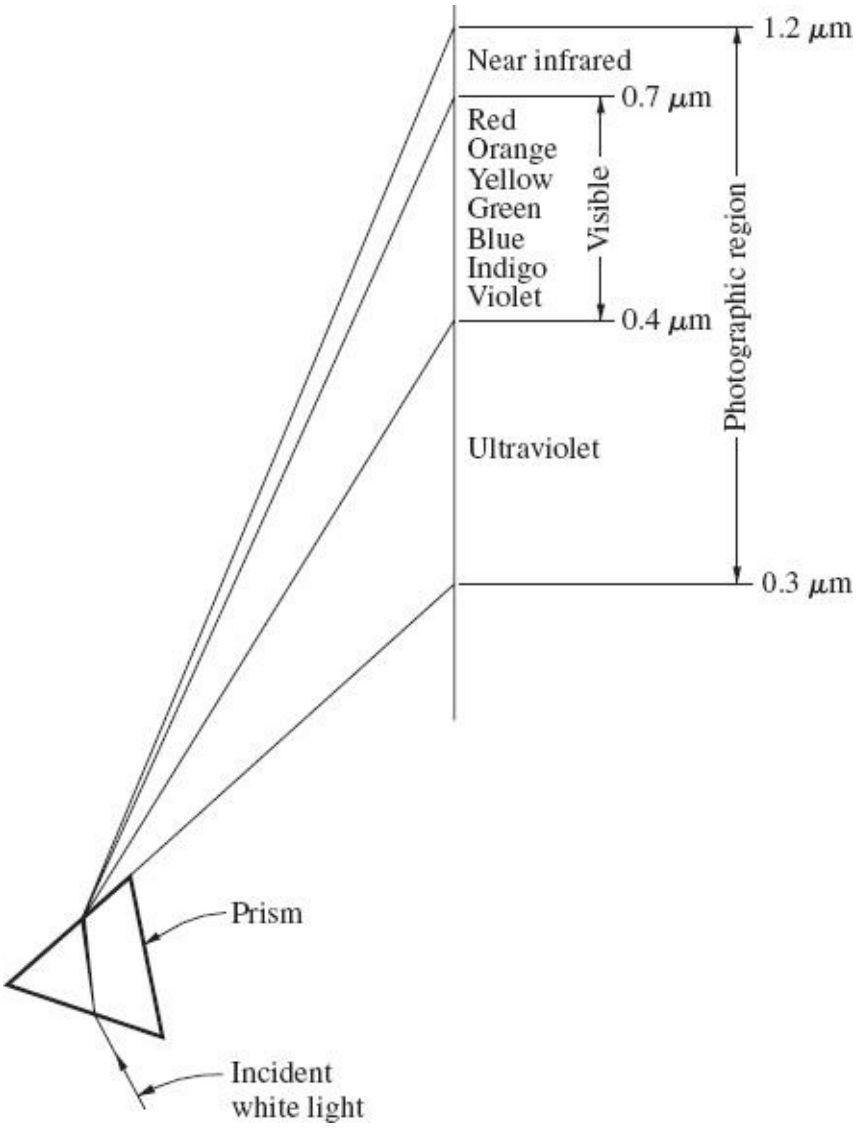


FIGURE 2-17 White light broken into the individual colors of the visible and near-visible spectrum by means of a prism. (Note that the range of wavelengths of transmitted light is non-linear.)

To the human eye, an object appears a certain color because the object reflects energy of the wavelengths producing that color. If an object reflects all the visible energy that strikes it, that object will appear white. But if an object absorbs all light and reflects none, that object will appear black. If an object absorbs all green and red energy but reflects blue, that object will appear blue.

Just as the retina of the human eye is sensitive to variations in wavelength, photographic emulsions can also be manufactured with variations in wavelength sensitivity. Black-and-white emulsions composed of untreated silver halides are sensitive only to blue and ultraviolet energy. Reflected light from a red object, for example, will not produce an image on such an emulsion. These untreated emulsions are usually used on printing papers for making positives from negatives. When these printing papers are used, red or yellow lights called *safe lights* can conveniently be used to illuminate the darkroom because these colors cannot expose a paper that is sensitive only to blue light.

Black-and-white silver halide emulsions can be treated by use of fluorescent dyes so that they are sensitive to other wavelengths of the spectrum besides blue. Emulsions sensitive to blue, green, and red are called *panchromatic*. Emulsions can also be made to respond to energy in the near-infrared range. These emulsions are called *infrared*, or IR. Infrared films make it possible to obtain photographs of energy that is invisible to the human eye. An early application of this type of emulsion was in camouflage detection, where it was found that dead foliage or green netting, which had the

same green color as live foliage to the human eye, reflected infrared energy differently. This difference could be detected through infrared photography. Infrared film is now widely used for a variety of applications such as detection of crop stress, tree species mapping, etc. Figure 2-18 illustrates sensitivity differences of various emulsions.

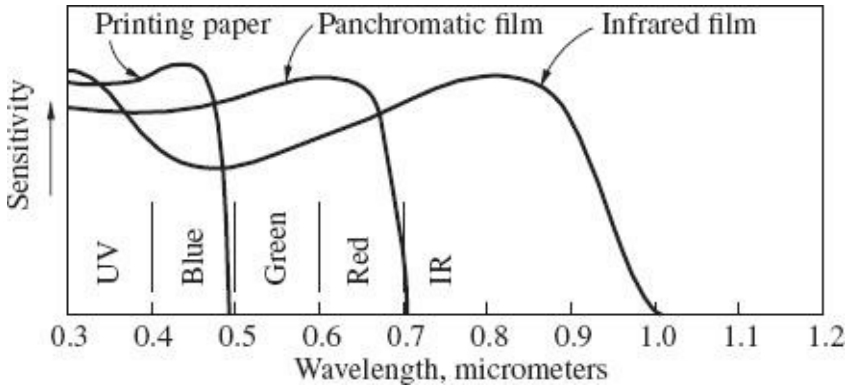


FIGURE 2-18 Typical sensitivities of various black-and-white emulsions.

2-10 Filters

The red or yellow safe light described in the previous section usually is simply an ordinary white light covered with a red or yellow filter. If the filter is red, it blocks passage of blue and green wavelengths and allows only red to pass. Filters placed in front of camera lenses also allow only certain wavelengths of energy to pass through the lens and expose the film. The use of filters on cameras can be very advantageous for certain types of photography.

Atmospheric haze is largely caused by the scattering of ultraviolet and short blue wavelengths. Pictures which are clear in spite of atmospheric haze can be taken through haze filters. These filters block passage of objectionable scattered short wavelengths (which produce haze) and prevent them from entering the camera and exposing the film. Because of this advantage, haze filters are almost always used on aerial cameras.

Filters for aerial mapping cameras are manufactured from high-quality optical glass. This is the case because light rays that form the image must pass through the filter before entering the camera. In passing through the filter, light rays are subjected to distortions caused by the filter. The camera should therefore be calibrated (see Secs. 3-10 through 3-14), with the filter locked firmly in place; after calibration, the filter should not be removed, for this would upset the calibration.

2-11 Color Film

Normal color and *color infrared* emulsions are used for certain applications in photogrammetry. Color emulsions consist of three layers of silver halides, as shown in Fig. 2-19. The top layer is sensitive to blue light, the second layer is sensitive to green and blue light, and the bottom layer is sensitive to red and blue light. A blue-blocking filter is built into the emulsion between the top two layers, thus preventing blue light from exposing the bottom two layers. The result is three layers sensitive to blue, green, and red light, respectively, from top to bottom. The sensitivity of each layer is indicated in Fig. 2-20.

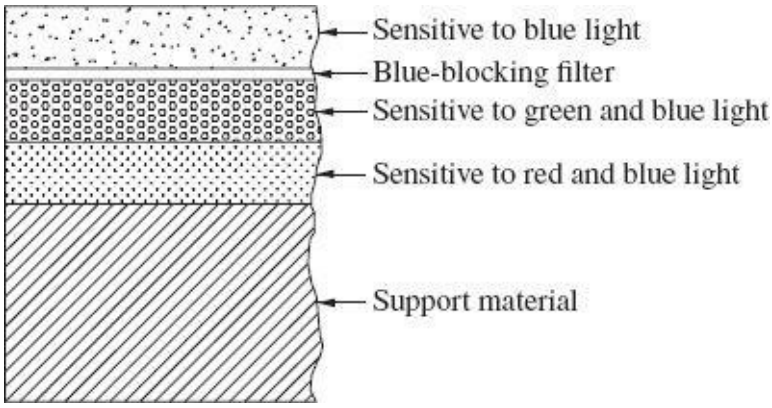


FIGURE 2-19 Cross section of normal color film.

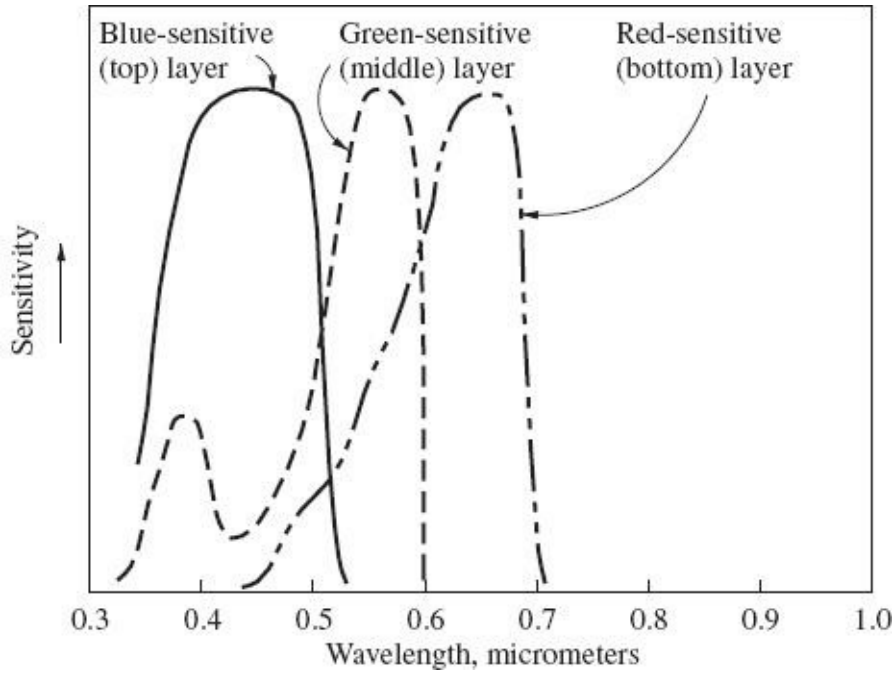


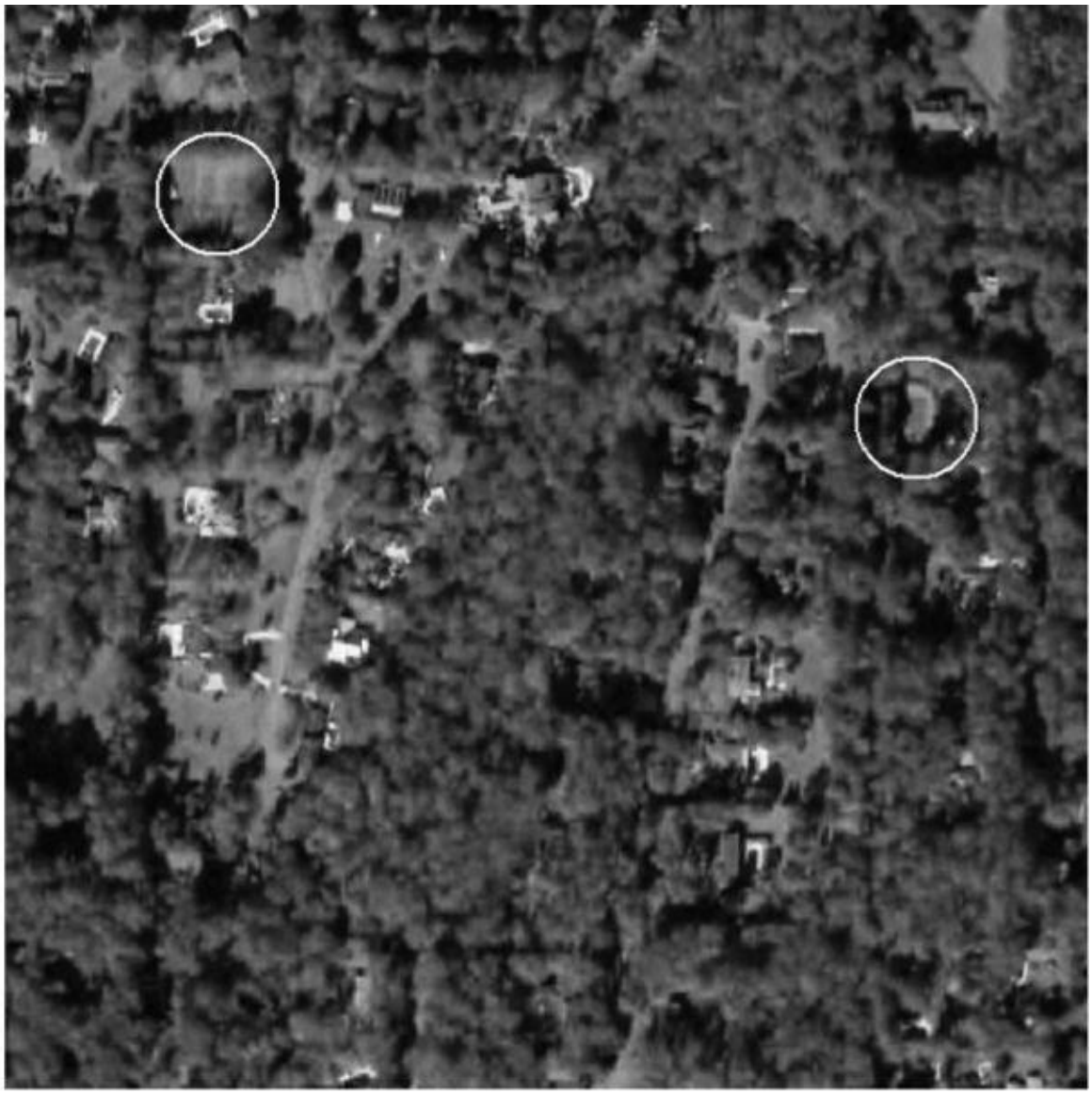
FIGURE 2-20 Typical color sensitivity of three layers of normal color film.

In making a color exposure, light entering the camera sensitizes the layer(s) of the emulsion that correspond(s) to the color or combination of colors of the original scene. There are a variety of color films available, each requiring a slightly different developing process. The first step of color developing accomplishes essentially the same result as the first step of black-and-white developing. The exposed halides in each layer are turned into black crystals of silver. The remainder of the process depends on whether the film is *color negative* or *color reversal* film. With color negative film, a negative is produced and color prints are made from the negative. Color reversal film produces a true color transparency directly on the film.

During World War II there was great interest in increasing the effectiveness of films in the infrared region of the spectrum. This interest led to the development of *color infrared* or *false-color* film. The military called it *camouflage detection* film because it allowed photo interpreters to easily differentiate between camouflage and natural foliage. Color Fig. 2-21*a* and *b* illustrate this effect. Green tennis courts with brown backgrounds can be seen within the circles of Fig. 2-21*a*, a normal color image. The color-infrared image of Fig. 2-21*b* depicts the same area, but uses various shades of red to represent reflected infrared energy. The tennis courts in Fig. 2-21*b* now appear with a grayish color, not red like that of the surrounding vegetation. Like normal color film, color IR film also has

three emulsion layers, each sensitive to a different part of the spectrum. Figure 2-22 illustrates the sensitivity curves for each layer of color IR film. The top layer is sensitive to ultraviolet, blue, and green energy. The middle layer has its sensitivity peak in the red portion of the spectrum, but it, too, is sensitive to ultraviolet and blue light. The bottom layer is sensitive to ultraviolet, blue, and infrared. Color IR film is commonly used with a yellow filter, which blocks wavelengths shorter than about 0.5 μm. The shaded area of Fig. 2-22 illustrates the blocking effect of a yellow filter.

(*a*)



(*b*)

FIGURE 2-21 (*a*) Normal color image and (*b*) color infrared image. Note that healthy vegetation, which appears green in the normal color image, appears red in the color infrared image. Circled tennis courts are painted green but appear gray in the color infrared image. (See also color insert.)

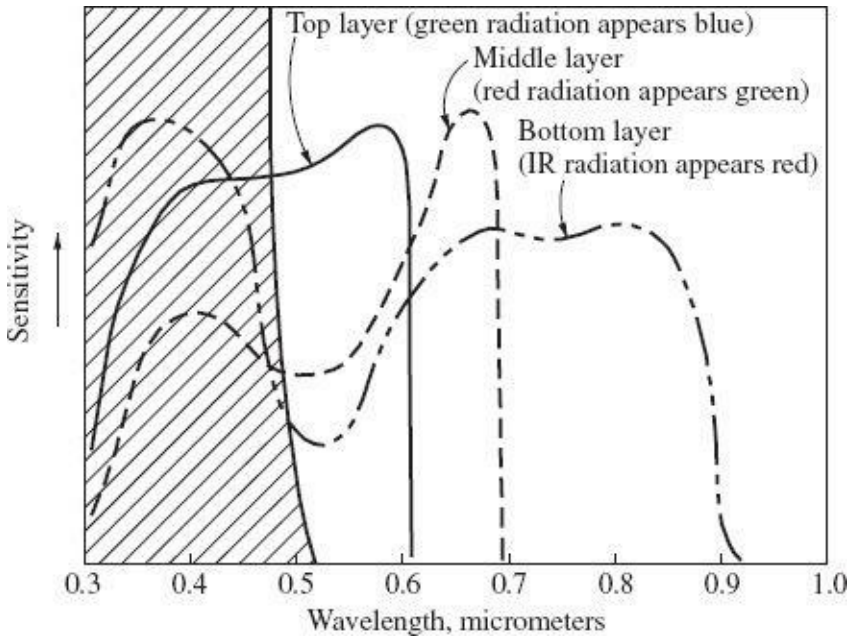
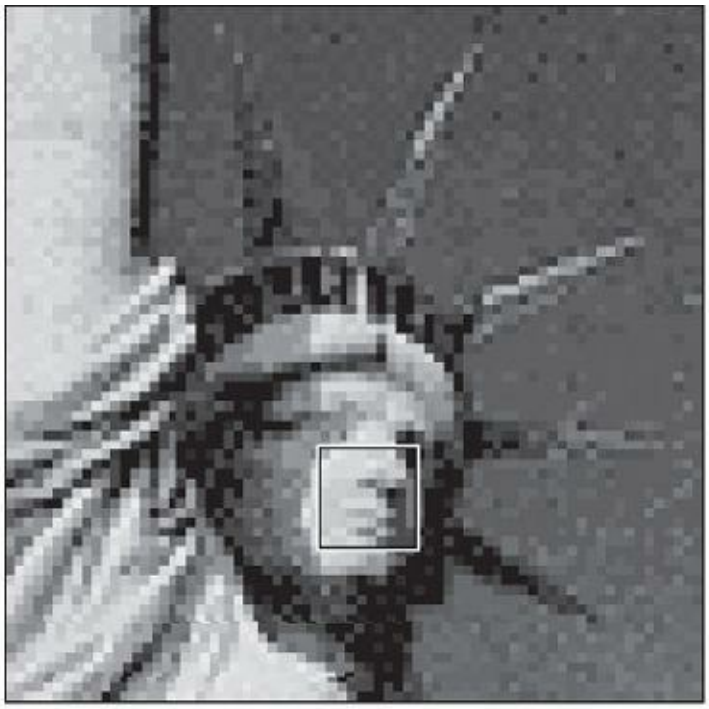


FIGURE 2-22 Typical sensitivity of color infrared (false-color) film.

To view the exposure resulting from invisible IR energy, the IR sensitive layer must be represented by one of the three primary (visible) colors. With color IR film and a yellow filter, any objects that reflect infrared energy appear red on the final processed picture. Objects that reflect red energy appear green, and objects reflecting green energy appear blue. It is this misrepresentation of color which accounts for the name *false color*. Although color IR film was developed by the military, it has found a multitude of uses in civilian applications.

2-12 Digital Images

A digital image is a computer-compatible pictorial rendition in which the image is divided into a fine grid of "picture elements," or *pixels*. The image in fact consists of an array of integers, often referred to as *digital numbers*, each quantifying the *gray level*, or degree of darkness, at a particular element. When an output image consisting of many thousands or millions of these pixels is viewed, the appearance is that of a continuous-tone picture. The image of the famous Statue of Liberty shown in Fig. 2-23*a* illustrates the principle. This image has been divided into a pixel grid of 72 rows and 72 columns, with each pixel being represented by a value from 0 (dark black) to 255 (bright white). A portion of the image near the mouth of the statue is shown enlarged in Figure 2-23*b*, and the digital numbers associated with this portion are listed in Fig. 2-23*c*. Note the correspondence of low numbers to dark areas and high numbers to light areas.



(a)

(b)

190	237	234	223	227	220	219	231	115	2
237	227	223	228	229	237	229	219	190	1
231	227	223	227	229	219	196	216	217	96
229	218	220	219	160	120	164	183	127	136
219	218	219	213	214	210	113	2	54	127
217	213	223	227	223	222	199	54	70	128
219	217	207	196	183	187	207	149	74	126
217	216	210	218	217	203	145	70	73	127
207	223	227	203	145	127	200	136	75	80
227	219	218	223	219	190	115	70	71	74

(c)

FIGURE 2-23 Digital image of Statue of Liberty.

The seemingly unusual range of values (0 to 255) can be explained by examining how computers deal with numbers. Since computers operate directly in the binary number system, it is most efficient and convenient to use ranges corresponding to powers of 2. Numbers in the range 0 to 255 can be accommodated by 1 *byte*, which consists of 8 binary digits, or *bits*. An 8-bit value can store 2^8, or 256, values, which exactly matches the range of 0 to 255 (remember to count the 0 as a value). The entire image of Fig. 2-23*a* would require a total of 72 × 72 = 5184 bytes of computer storage.

Digital images are produced through a process referred to as *discrete sampling*. In this process, a small image area (a pixel) is "sensed" to determine the amount of electromagnetic energy given off by the corresponding patch of surface on the object. Discrete sampling of an image has two fundamental characteristics, *geometric resolution* and *radiometric resolution*. Geometric (or spatial) resolution refers to the physical size of an individual pixel, with smaller pixel sizes corresponding to higher geometric resolution. The four illustrations of Fig. 2-24 show the entire image of the Statue of Liberty and demonstrate the effect of different geometric resolutions on image clarity. The original 72 × 72

pixel image of Fig. 2-24*a* and the half-resolution 36 × 36 image of Fig. 2-24*b* are readily discernible. The 18 × 18 image of Fig. 2-24*c* is barely recognizable, and then only when the identity of the actual feature is known. At the resolution of the 9 × 9 image of Fig. 2-24*d*, one sees a semiorganized collection of blocks bearing little resemblance to the original image, although the rough position of the face and the arm can be detected. Obviously, geometric resolution is important for feature recognition in digital photographs.

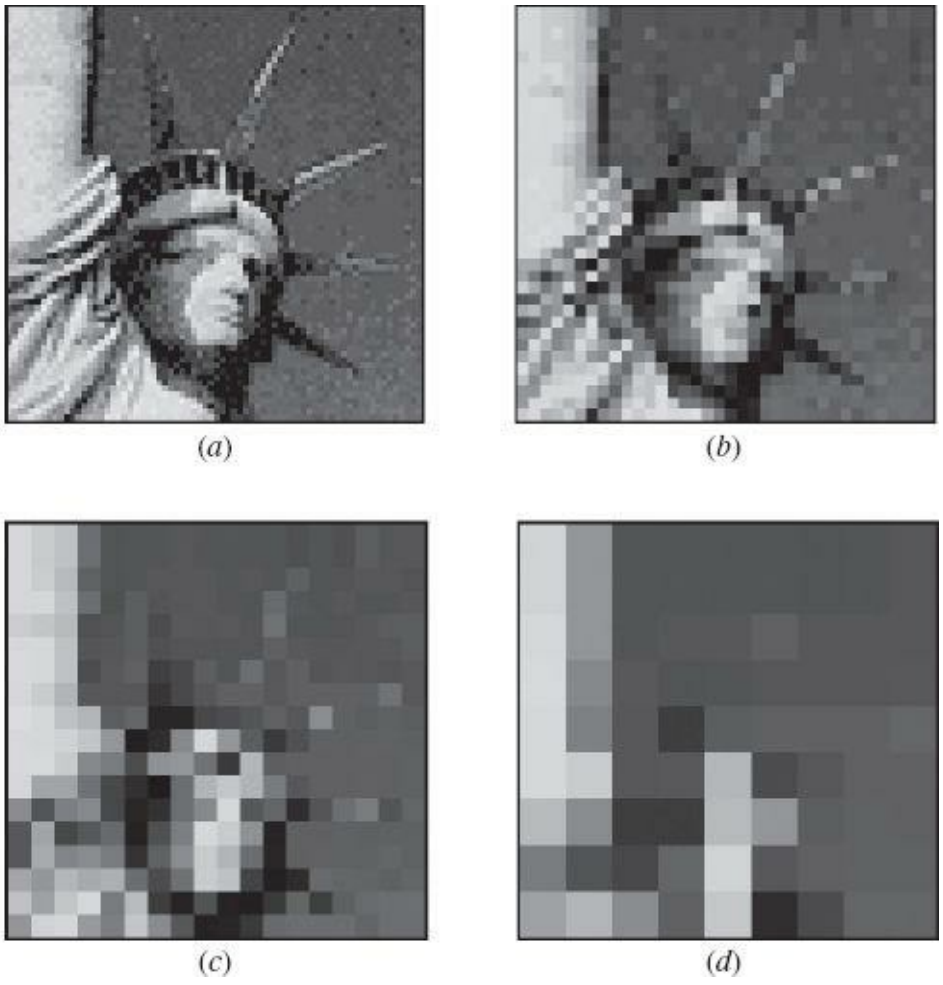


FIGURE 2-24 Digital image of Statue of Liberty at various spatial resolutions.

Another fundamental characteristic of digital imagery is radio-metric resolution, which can be further broken down into level of *quantization* and *spectral resolution*. Quantization refers to the conversion of the amplitude of the original electromagnetic energy (analog signal) into a number of discrete levels (digital signal). Greater levels of quantization result in more accurate digital representations of the analog signal. While the image of Fig. 2-23 has 8-bit quantization, some digital images have 10-bit (1024 gray levels), 12-bit (4096 gray levels), or even higher quantization.

Figure 2-25 illustrates the effect of various levels of quantization for the statue image. In this figure, part (*a*) shows the original image with 256 discrete quantization levels, while parts (*b*), (*c*), and (*d*) show 8, 4, and 2 levels, respectively. Notice that in lower-level quantizations, large areas appear homogeneous and subtle tonal variations can no longer be detected. At the extreme is the 2-level quantization, which is also referred to as a *binary* image. Quantization level has a direct impact on the amount of computer memory required to store the image. The binary image requires only 1 bit of computer memory for each pixel ($2^1 = 2$), the 4-level image requires 2 bits ($2^2 = 4$), and the 8-level image requires 3 bits ($2^3 = 8$).

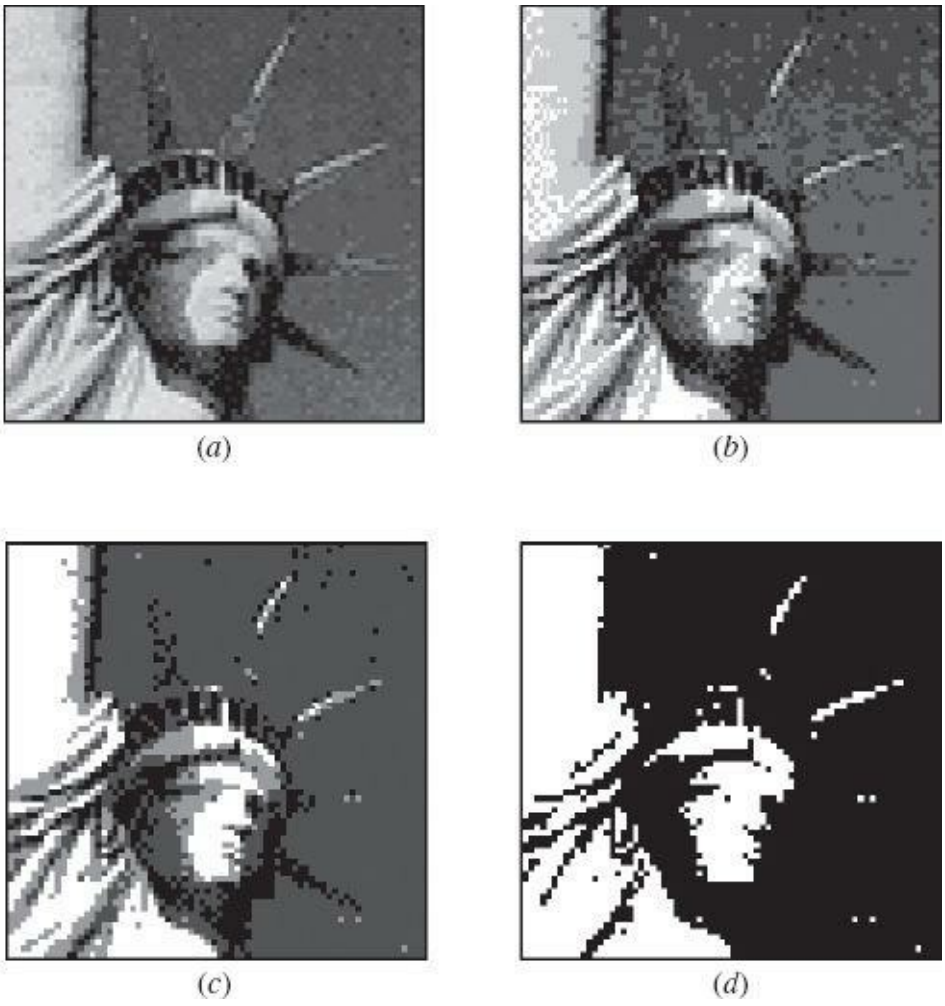


FIGURE 2-25 Digital image of Statue of Liberty at various quantization levels.

Spectral resolution is another aspect of radiometric resolution. Recall from Sec. 2-9 that the electromagnetic spectrum covers a continuous range of wavelengths. In imaging, samples of electromagnetic energy are normally sensed only within narrow bands in the spectrum. For example, the color film discussed in Sec. 2-9 samples three bands within the visible portion of the spectrum—blue, green, and red. The spectral resolution of color film corresponds to three spectral bands or channels. The black-and-white statue image of Fig. 2-23 has only one spectral band which covers the entire visible range from violet through red. Other imaging devices may take samples in several, or even hundreds, of bands. With a higher level of spectral resolution, a more accurate representation of an object's *spectral response pattern* can be made. Since each spectral band must be stored individually in computer memory, the number of bands has a direct impact on the amount of image storage required.

Example 2-4
A 3000-row by 3000-column satellite image has three spectral channels. If each pixel is represented by 8 bits (1 byte) per channel, how many bytes of computer memory are required to store the image?

Solution

$$\text{Number of pixels} = 3000 \times 3000 = 9{,}000{,}000 \text{ pixels}$$

$$\text{Bytes per pixel} = (3 \text{ channels}) \times (1 \text{ byte}) = 3 \text{ bytes}$$

$$\text{Image size} = (9{,}000{,}000 \text{ pixels}) \times (3 \text{ bytes/pixel})$$

$$= 27{,}000{,}000 \text{ bytes} \quad \blacktriangle$$

2-13 Color Image Representation

A human's perception of color is a complex interaction involving an analog spectral signal. Highly accurate and precise portrayal of the analog signal in the digital realm is prohibitively expensive in terms of memory storage. Color digital images generally utilize a simplification based on the same principle as a television screen. With the set turned on, hold a magnifying lens up to a white area on your TV (or computer) screen. You should see that the region is covered by thousands of tiny red, green, and blue dots or rectangles. If you look at an area of the screen that is a bluish color, you will see that the blue dots remain brightly lit, whereas the red and green dots are dim. At normal viewing distance, however, your eyes perceive seemingly continuous shades of color. In this same fashion, digital images represent shades of color by varying the levels of brightness of the three primary colors or channels, individually. The samples from these three channels form the digital representation of the original analog signal. This scheme is appropriately based on the theory that three types of cones (color-sensing components of the retina) in the human eye are likewise most sensitive to blue, green, and red energy.

The color of any pixel can be represented by three-dimensional coordinates in what is known as BGR (or RGB) color space. Assume for example that the digital number in each channel is quantified as an 8-bit value, which can range from 0 to 255. The color of a pixel is thus represented by an ordered triplet consisting of a blue value, a green value, and a red value, each of which can range from 0 to 255. This can be represented by the three-dimensional axis system shown in Fig. 2-26. This figure shows the *color cube*, where a particular color has a unique set of *B*, *G*, *R* coordinates. Figure 2-27 is a color illustration of the color cube from (*a*) behind the origin and (*b*) from the opposite corner. Note how the colors change from one region of the cube to another.

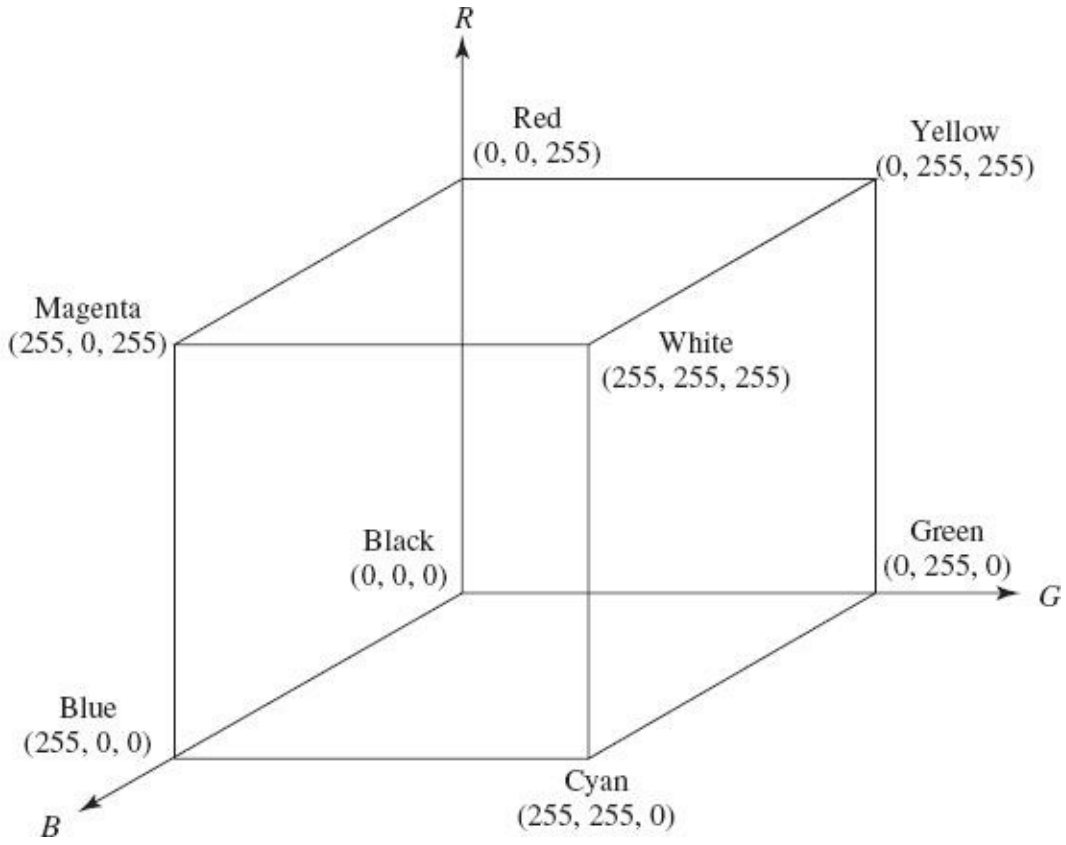


FIGURE 2-26 Blue-green-red color cube.

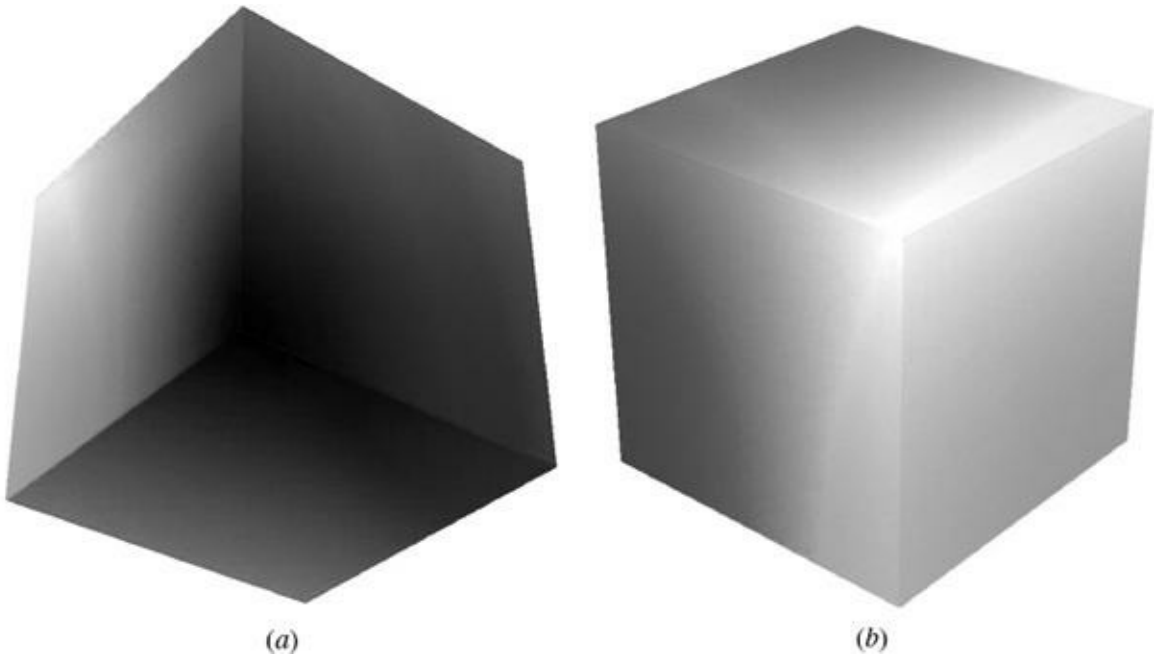


FIGURE 2-27 (*a*) A view of the color cube from behind the origin and (*b*) a view of the color cube from the opposite corner. (See also color insert.)

Even though a *B,G,R* position within the color cube is adequate to specify any given color, this scheme does not lend itself to convenient human interpretation. The *intensity-hue-saturation* (IHS) system, on the other hand, is more readily understood by humans. This system can be defined as a set of cylindrical coordinates in which the height, angle, and radius represent intensity, hue, and saturation, respectively. The axis of the cylinder is the *gray line* which extends from the origin of the color cube to the opposite corner where the levels of blue, green, and red are maximum, as shown in

Fig. 2-28. Thus, for any particular color, intensity represents overall brightness irrespective of color, hue represents the specific mixture of wavelengths that define the color, and saturation represents the boldness of the color. The representations of hue and saturation are illustrated in Fig. 2-29, which is a two-dimensional view showing the projection of the color cube and base of the cylinder as viewed along the gray line toward the origin. Here, hue and saturation appear as a set of polar coordinates with the direction of the 0° hue axis being arbitrarily chosen as halfway between blue and magenta. As an example, in Fig. 2-29 the hue value of 240° (–120°) corresponds to the color orange, and the large saturation value (radius) represents a very bold or vivid color. Figure 2-30 is a color image that shows the color cube in a position corresponding to Fig. 2-29. This figure allows a visual interpretation of the effects of hue and saturation.

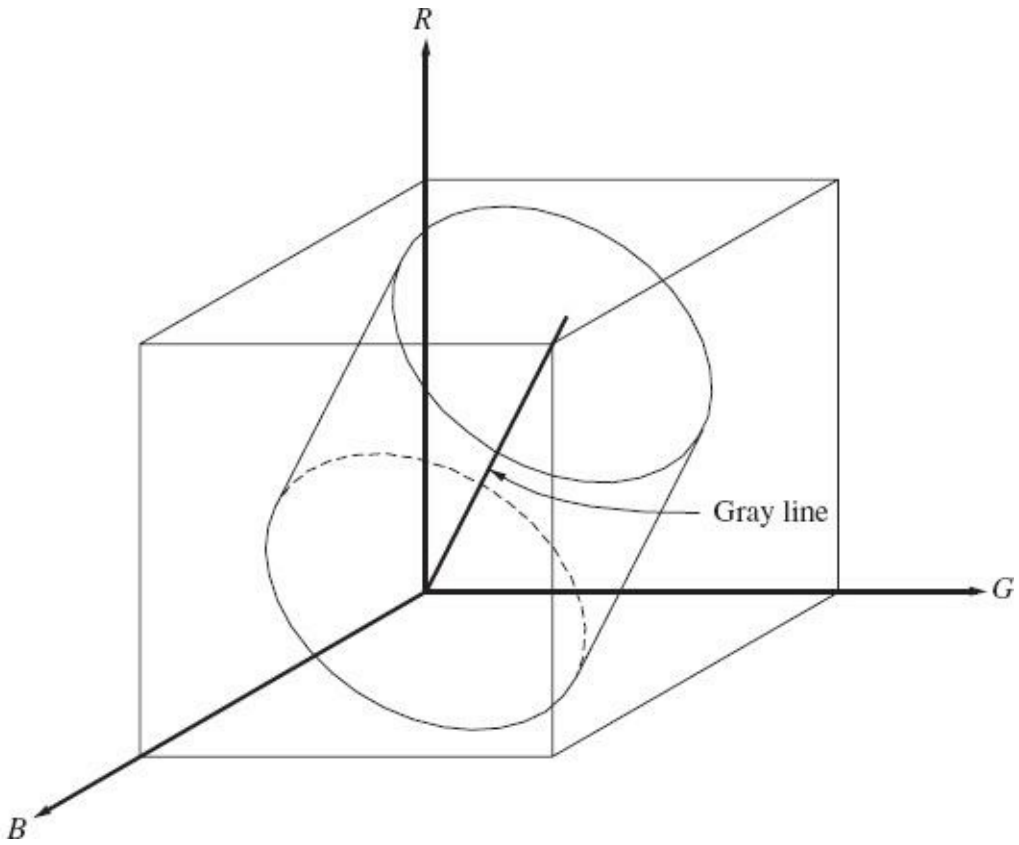


FIGURE 2-28 Position of gray line as axis of cylindrical intensity-hue-saturation system.

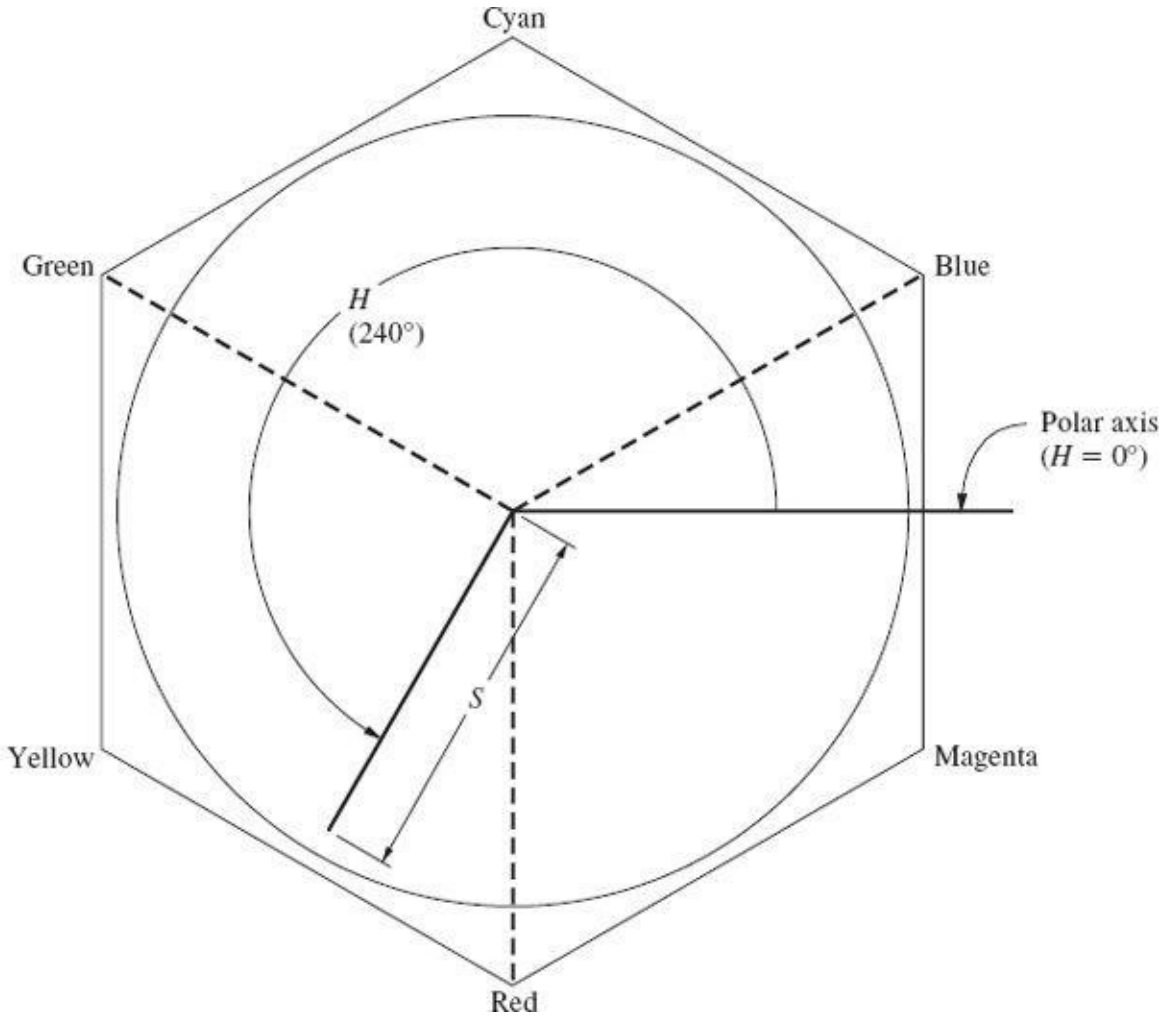


FIGURE 2-29 Representation of hue and saturation with respect to color values.

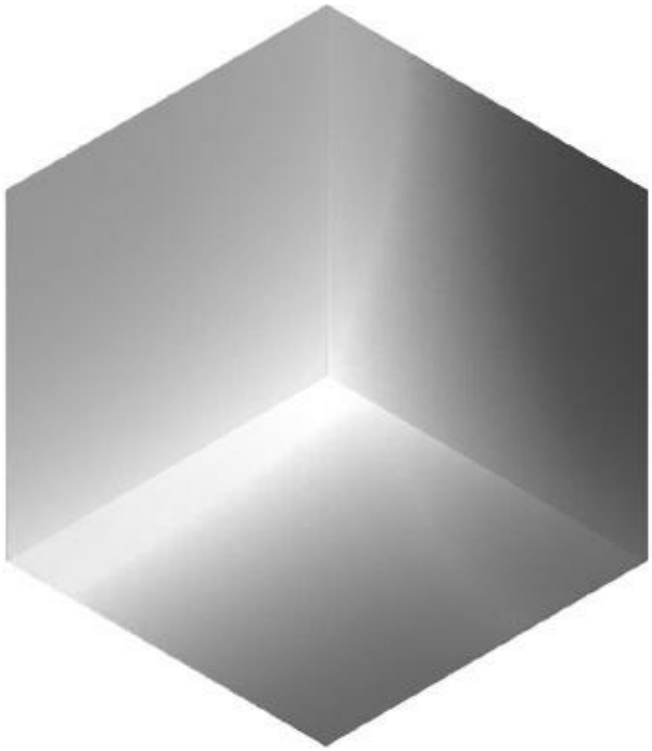


FIGURE 2-30 Representation of hue and saturation corresponding to Fig. 2-29. (See also color insert.)

Conversion from the system of *B*, *G*, *R* coordinates to *I*, *H*, *S* is accomplished by starting with the

intensity (cylinder) axis lined up with the red axis and then rotating so the intensity axis lines up with the gray line. Then the standard cartesian-to-cylindrical coordinate conversion is applied. To convert from *I*, *H*, *S* to *B*, *G*, *R* the process is reversed. The derivation is based on a rigorous three-dimensional coordinate conversion, the details of which are not presented in this text. To convert from *B*, *G*, *R* to *I*, *H*, *S*, first compute two intermediate variables *X* and *Y* by Eqs. (2-8) and (2-9). Then intensity, hue, and saturation can be computed by Eqs. (2-10), (2-11), and (2-12), respectively. Example 2-5 demonstrates the procedure.

$$X = \frac{B - G}{\sqrt{2}} \tag{2-8}$$

$$Y = \frac{B + G - 2R}{\sqrt{6}} \tag{2-9}$$

$$I = \frac{B + G + R}{\sqrt{3}} \tag{2-10}$$

$$H = \tan^{-1}\left(\frac{Y}{X}\right) \tag{2-11}$$

$$S = \sqrt{X^2 + Y^2} \tag{2-12}$$

In order for the full range of hues to be accommodated, the full-circle inverse tangent function (e.g., ATAN2 in many programming languages or spreadsheets) must be used in Eq. (2-11). To convert back from *I*, *H*, *S* to *B*, *G*, *R* intermediate values of *X* and *Y* are computed by Eqs. (2-13) and (2-14). Then blue, green, and red can be computed by Eqs. (2-15), (2-16), and (2-17), respectively.

$$X = S\cos H \tag{2-13}$$

$$Y = S\sin H \tag{2-14}$$

$$B = \frac{X}{\sqrt{2}} + \frac{Y}{\sqrt{6}} + \frac{I}{\sqrt{3}} \tag{2-15}$$

$$G = -\frac{X}{\sqrt{2}} + \frac{Y}{\sqrt{6}} + \frac{I}{\sqrt{3}} \tag{2-16}$$

$$R = -\frac{2Y}{\sqrt{6}} + \frac{I}{\sqrt{3}} \tag{2-17}$$

When conversion is done from *B, G, R* to *I, H, S* the resulting value of intensity can range from 0 to $255\sqrt{3}$, hue can range from $-\pi$ to $+\pi$, and saturation can range from 0 to $255\sqrt{\frac{2}{3}}$, (assuming the ranges for the *B, G, R* coordinates were each 0 to 255). Since it is often desirable to store the *I, H, S* values as 1-byte integers, they may have to be rescaled to the 0-to-255 range. When these values are rescaled, the conversion process may no longer be perfectly reversible due to loss of precision. In other words, given a particular set of *B, G, R* coordinates, if they are converted to *I, H, S* scaled to a 1-byte integer from 0 to 255, rescaled to the original ranges, and converted back to *B, G, R* then the final values may be slightly different from the original ones.

Example 2-5
Convert the *B, G, R* coordinates (165, 57, 105) to *I, H, S* values. Then convert the resultant values back to *B, G, R*.

Solution By Eqs. (2-8) through (2-12):

$$X = \frac{165-57}{\sqrt{2}} = 76.37$$

$$Y = \frac{165+57-2\times 105}{\sqrt{6}} = 4.899$$

$$I = \frac{165+57+105}{\sqrt{3}} = 188.8$$

$$H = \tan^{-1}\left(\frac{Y}{X}\right) = 0.06406 \text{ rad} = 3.670^\circ$$

$$S = \sqrt{X^2+Y^2} = 76.52$$

Reverse by using Eqs. (2-13) through (2-17):

$$X = 76.52\ \cos(0.06406 \text{ rad}) = 76.36$$

$$Y = 76.52\ \sin(0.06406 \text{ rad}) = 4.899$$

$$B = \frac{76.36}{\sqrt{2}} + \frac{4.899}{\sqrt{6}} + \frac{188.8}{\sqrt{3}} = 165$$

$$G = -\frac{X}{\sqrt{2}} + \frac{Y}{\sqrt{6}} + \frac{I}{\sqrt{3}} = 57$$

$$R = -\frac{2Y}{\sqrt{6}} + \frac{I}{\sqrt{3}} = 105 \quad \blacktriangle$$

2-14 Digital Image Display

The *B, G, R* system discussed in the preceding section is an example of a *color-additive* process. In that process, different colors of actively produced light are combined to yield other colors. The color-

additive process is used to produce images on typical TV or computer displays. When a color digital image is printed on paper, the situation is different. Instead of each pixel being an active light source, ambient white light reflects off colored inks, which absorb selected colors or channels of energy. Digital printing is a *color-subtractive* process because certain colors are absorbed (subtracted) from the ambient white light. Figure 2-31*a* illustrates the mixing of colors in a color additive process. The three circles represent active light sources of red, green, and blue light. The regions of intersection show the resulting colors that occur when the light sources are combined. Red and green light combine to produce yellow, green and blue produce cyan, and blue and red produce magenta. Figure 2-31*b* illustrates the mixing of colors in a color subtractive process. The three circles represent pigments or inks having the colors of cyan, magenta, and yellow. The regions of intersection show the resulting colors that occur when the pigments or inks are combined. Cyan and magenta combine to produce blue, magenta and yellow produce red, and yellow and cyan produce green.

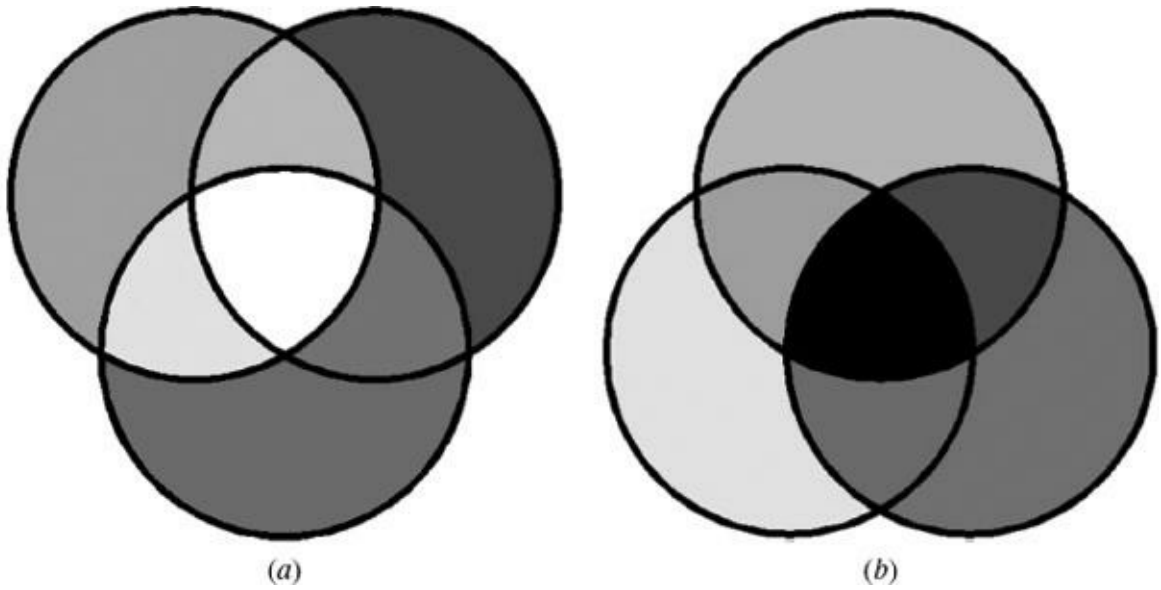


FIGURE 2-31 Illustration of the (*a*) color additive process and (*b*) color subtractive process. (See also color insert.)

The display of a digital color image on TV or computer monitor is straightforward. The area within each pixel emits light of various intensities for each of the primary colors (red, green, and blue), which mix to produce the appropriate color. Printing of color images uses the process of color subtraction, but it is more complicated because the intensity of pigment or ink cannot be easily controlled. The *complementary* colors of yellow, magenta, and cyan are used instead of the primary colors of blue, green, and red. For example, yellow ink appears yellow because the blue portion of the incident white light is absorbed by the ink, and only the green and red energy are reflected. Similarly, magenta absorbs green energy, reflecting only blue and red, and cyan absorbs red energy, reflecting only blue and green. To produce the color red, yellow ink and magenta ink are used, which will absorb blue and green, respectively, leaving a red hue. By varying the spatial density of the ink dots, the amounts of blue, green, and red energy being absorbed will likewise vary, allowing the full range of colors to be produced. Generally, black ink is incorporated into the scheme, which allows darker grays and black to be produced.

References

American Society of Photogrammetry: *Manual of Color Photography*, Bethesda, MD, 1968.
_____: *Manual of Photogrammetry*, 4th ed., Bethesda, MD, 1980, chaps. 3 and 6.
American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004, Chap. 4.
Baines, H., and E. S. Bomback: *The Science of Photography*, Halsted Press, New York, 1974.
Carlson, F. P.: *Introduction to Applied Optics for Engineers*, Academic Press, New York, 1972.
Fent, L., R. J. Hall, and R. K. Nesby: "Aerial Films for Forest Inventory: Optimizing Film Parameters," *Photogrammetric Engineering and Remote Sensing*, vol. 61, no. 3, 1995, p. 281.
Ghatak, A.: *An Introduction to Modern Optics*, McGraw-Hill, New York, 1972.
Handbook of Chemistry and Physics, 56th edition, CRC Press, Cleveland, OH, 1975.
Harris, J. R., R. Murray, and T. Hirose: "IHS Transform for the Integration of Radar Imagery with Other Remotely Sensed Data," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 12, 1990, p. 1631.
Klimes, D., and D. I. Ross: "A Continuous Process for the Development of Kodak Aerochrome Infrared Film 2433 as a Negative," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 2, 1993, p. 209.
Knuth, D. E.: "Digital Halftones by Dot Diffusion," *ACM Transactions on Graphics*, vol. 6, no. 4, 1987, p. 245.
Laikin, M.: *Lens Design*, 2d ed., Marcel Dekker, New York, 1995.
Lillesand, T. M., R. W. Kiefer, and J. W. Chipman: *Remote Sensing and Image Interpretation*, 6th ed., Wiley, New York, 2008.
Malacara, D., and Z. Malacara: *Handbook of Lens Design*, Marcel Dekker, New York, 1994.
Smith, W. J., and Genesee Optics Software, Inc.: *Modern Lens Design—A Resource Manual*, McGraw-Hill, New York, 1992.
Tani, T.: *Photographic Sensitivity—Theory and Mechanisms*, Oxford University Press, New York, 1995.
Walther, A.: *The Ray and Wave Theory of Lenses*, Cambridge University Press, New York, 1995.

Problems

2-1. Briefly explain the difference between physical and geometric optics.

2-2. Under certain conditions, the speed of light through air is 2.99688×10^8 m/s. What is the index of refraction of air under these conditions?

2-3. A certain electromagnetic energy is propagated in a vacuum at a frequency of 2,361,000 cycles per second. What is the wavelength (to the nearest meter) of this energy?

2-4. The wavelength of visible light ranges from 0.4 to 0.7 μm. Express this range in terms of frequency (to the nearest cycle per second), based on the speed of light in a vacuum.

2-5. If a certain type of glass has an index of refraction of 1.550, what is the speed of light through this glass?

2-6. A ray of light enters glass (index 1.570) from air (index 1.0003) at an incident angle of 36.9°. Find the angle of refraction.

2-7. A ray of light enters glass (index 1.515) from air (index 1.0003) at an incident angle of 67.8°. Find the angle of refraction.

2-8. A light ray emanating from under water (index 1.333) makes an angle of 37.5° with the normal to the surface. What is the angle that the refracted ray makes with the normal as it emerges into air (index 1.0003)?

2-9. What is the incident angle for a light ray emanating from glass (index 1.53) as it passes into air (index 1.0003) so that the angle of refraction is 90.0°? (This angle is known as the *critical angle*, above which total internal reflection takes place.)

2-10. Repeat Prob. 2-9, except the light ray emanates from water (index 1.333) into air.

2-11. A single ray of light traveling through air (index 1.0003) enters a convex glass lens (index 1.575) having a radius of 47.5 mm. If the light ray is parallel to and 9.5 mm above the optical axis of the lens, what are the angles of incidence and refraction?

2-12. An object located 1.8 m in front of a thin lens has its image in focus 72.5 mm from the lens on the other side. What is the focal length of the lens?

2-13. An object is located 12 m in front of a thin lens having a focal length of 50.0 mm. At what image distance will the object's image be in perfect focus?

2-14. A camera lens can accommodate object distances ranging from 1.2 m to infinity. If the focal length of the lens is 38 mm, what is the corresponding range of image distances?

2-15. A lens has a focal length of 70.0 mm. What is the object distance for an image that is perfectly focused at an image distance of 125.0 mm?

2-16. Prepare a table of image distances (in millimeters) versus object distances of exactly 1, 2, 5, 10, 100, 1000, and 5000 m for a lens having a 152.416-mm focal length, such that the images are in perfect focus.

2-17. Explain why the lens camera replaced the early pinhole camera.

2-18. Define the photographic terms *illuminance, aperture, emulsion, latent image,* and *fog*.

2-19. A camera lens has a focal length of 35.0 mm. Its *f*-stop settings range from *f*-1.4 to *f*-22. What is the maximum diameter of the aperture? Minimum diameter?

2-20. A camera with a 50.0-mm focal length lens has the *f*-stop set at 4. A 50.0-mm cylindrical extension is inserted between the lens and the camera body, increasing the nominal image distance from 50 to 100-mm. What true *f*-stop corresponds to the original setting of 4?

2-21. Prepare a table of lens aperture diameters versus nominal *f*-stop settings ranging from *f*-1 to *f*-32 for a 70.0-mm focal length lens.

2-22. An exposure is optimum at a shutter speed of $\frac{1}{250}$ s and *f*-8. If it is necessary to change the shutter speed to $\frac{1}{500}$ s, what should be the corresponding *f*-stop, to retain optimum exposure?

2-23. Repeat Prob. 2-22, except change the shutter speed to $\frac{1}{125}$ s.

2-24. An exposure is optimum at a shutter speed of $\frac{1}{1000}$ s and *f*-5.6. To increase the depth of field, it is necessary to expose at *f*-22. What is the required shutter speed to retain optimum exposure?

2-25. What is the relationship between film speed and emulsion grain size?

2-26. What is the relationship between resolution and emulsion grain size?

2-27. Explain how the slope of the straight-line portion of the *D*–log *E* curve relates to contrast.

2-28. Explain when and why a safe light can be used in a darkroom.

2-29. Explain why a haze filter is used on aerial cameras.

2-30. A panchromatic (one-band) digital image has 16,000 rows and 16,000 columns. Assuming each pixel takes 1 byte, how many bytes are required to store the image?

2-31. A 6200-row by 6200-column digital image has six spectral channels. Assuming each pixel requires 1 byte per channel, how many bytes are required to store the image?

2-32. Convert the *B, G, R* color of (115, 55, 135) to the *I, H, S* representation.

2-33. Convert the *I, H, S* color of (221, 225°, 65) to the *B, G, R* representation.

2-34. Describe how the color subtractive process can be applied to produce green.

CHAPTER 3

Cameras and Other Imaging Devices

3-1 Introduction

Perhaps the most fundamental device in the field of photogrammetry is the camera. It is the basic instrument which acquires images, from which photogrammetric products are produced.

The fourth edition of the *Manual of Photogrammetry* defines a *camera* as "a lightproof chamber or box in which the image of an exterior object is projected upon a sensitized plate or film, through an opening usually equipped with a lens or lenses, shutter, and variable aperture." That definition has been broadened in recent years with the increased use of the *digital camera* which senses light energy through the use of semiconductor electronics instead of film. In many cases a more general term such a s *imaging device* may be more appropriate to describe the instrument used for primary photogrammetric data acquisition.

Whatever the characteristics of the imaging device may be, an understanding of the underlying geometry is essential for precise and accurate applications of photogrammetry. The remarkable success of photogrammetry in recent years is due in large part to the progress that has been made in developing precision cameras. Perhaps the most noteworthy among recent camera developments has been the perfection of lenses of extremely high resolving power and almost negligible distortion. This has greatly increased the accuracy of photogrammetry. There have also been many significant improvements in general camera construction and operation.

Imaging devices can be categorized according to how the image is formed. Devices that acquire the image simultaneously over the entire format are *frame cameras* (or *frame sensors*). Frame cameras generally employ shutters which open and allow light from the field of view to illuminate a two-dimensional (usually rectangular) image plane before closing. Other imaging devices sense only a linear projection (strip) of the field of view at a given time and require that the device move or sweep across the area being photographed in order to acquire a two-dimensional image. Devices of this second type are referred to as *strip cameras, linear array sensors,* or *pushbroom scanners*. A third type of device builds an image by detecting only a small spot at a time, requiring movements in two directions (sweep and scan) in order for the two-dimensional image to be formed. These devices are often referred to as *flying spot scanners* or *whiskbroom scanners*.

The traditional imaging device used in photogrammetry is the aerial mapping camera, and its use is widespread in the photogrammetric industry. The requirements of aerial mapping cameras are quite different from those of ordinary handheld cameras, such as that shown in Fig. 2-11. The primary requirement of any photogrammetric aerial camera is a lens of high geometric quality. Aerial cameras must be capable of exposing in rapid succession a great number of photographs to exacting specifications. Since these cameras must perform this function while moving in an aircraft at high speed, they must have short cycling times, fast lenses, and efficient shutters. They must be capable of faithful functioning under the most extreme weather conditions and in spite of aircraft vibrations. Aerial cameras using roll film generally have magazine capacities of several hundred exposures while digital mapping cameras typically have sufficient computer memory to store equivalent numbers of images. Because the aerial photographic flight mission is fairly expensive and since weather and other conditions may prevent aerial photography for long periods of time, it is imperative that every precaution be taken in the manufacture of aerial cameras to guarantee the quality and reliability of the

photography on each mission.

This chapter discusses various types of imaging devices, but the standard-film aerial mapping camera is presented in the greatest detail. This is primarily due to its continued wide usage; however, this is also a practical approach because other imaging devices can then conveniently be described as variations of this basic instrument.

3-2 Metric Cameras for Aerial Mapping

Single-lens frame cameras are the most common cameras in use today. They are used almost exclusively in obtaining photographs for mapping purposes because they provide the highest geometric picture quality. With a single-lens frame camera, the lens is held fixed relative to the focal plane. The film is generally fixed in position during exposure, although it may be advanced slightly during exposure to compensate for image motion. The entire format is exposed simultaneously with a single click of the shutter.

Single-lens frame cameras are often classified according to their *angular field of view*. Angular field of view, as illustrated in Fig. 3-1, is the angle α subtended at the rear nodal point of the camera lens by the diagonal d of the picture format. [The most common frame or format size of film aerial mapping cameras is 230 mm (9 in) square.] Classifications according to angular field of view are

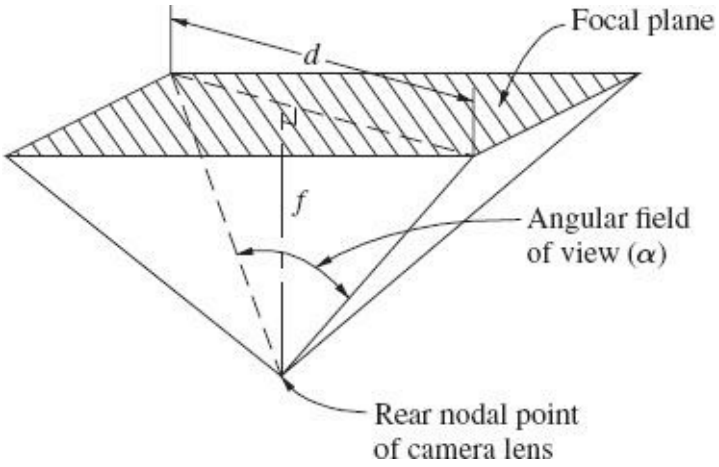


FIGURE 3-1 Angular field of view of a camera.

1. Normal angle (up to 75°)
2. Wide angle (75° to 100°)
3. Superwide angle (greater than 100°)

Angular field of view may be calculated as follows (see Fig. 3-1):

$$\alpha = 2\tan^{-1}\left(\frac{d}{2f}\right) \tag{3-1}$$

For a nominal 152-mm-focal-length camera with a 230-mm-square format, the angular field of view is

$$\alpha = 2\tan^{-1}\left[\frac{\sqrt{230^2+230^2}}{2(152)}\right] = 94^\circ \qquad \text{wide angle}$$

Single-lens frame cameras are available in a variety of lens *focal lengths*, and the choice depends on the purpose of the photography. The most common one in use today for mapping photography has a 152-mm (6-in) focal length and 230-mm- (9-in-) square format, although 89-mm ($3\frac{1}{2}$-in), 210-mm ($8\frac{1}{4}$-in), and 305-mm (12-in) focal lengths with 230-mm formats are also used. The 152-mm focal length with a 230-mm format provides the best combination of geometric strength and photographic scale for mapping. Longer focal lengths such as 305 mm are used primarily for obtaining photographs for aerial mosaics and for reconnaissance and interpretation purposes. They enable reasonably large photographic scales to be obtained in spite of high flying heights, and they reduce image displacements due to relief variations (see Sec. 6-8).

Digital mapping cameras come in a variety of formats and focal lengths, although most are designed to capture images in the normal angle range—narrower than those of the 152-mm focal length film camera.

From Eq. (3-1), it is seen that for a particular format size, the angular field of view increases as focal length decreases. Short focal lengths, therefore, yield wider ground coverage at a given flying height than longer focal lengths.

In Fig. 1-1 the Zeiss RMK TOP 15 aerial mapping camera was shown, and Fig. 3-2 illustrates the Leica RC30 aerial mapping camera. These cameras and their predecessors, together with a few others, are being used today to take the bulk of aerial-film photography for mapping purposes. Both are precision single-lens frame cameras having 230-mm-square formats and film capacities of approximately 500 exposures. The TOP 15 has a nominal 152-mm-focal-length lens. The RC30 is capable of accepting interchangeable cones with lenses having nominal focal lengths of 89, 152, 210, or 305 mm.

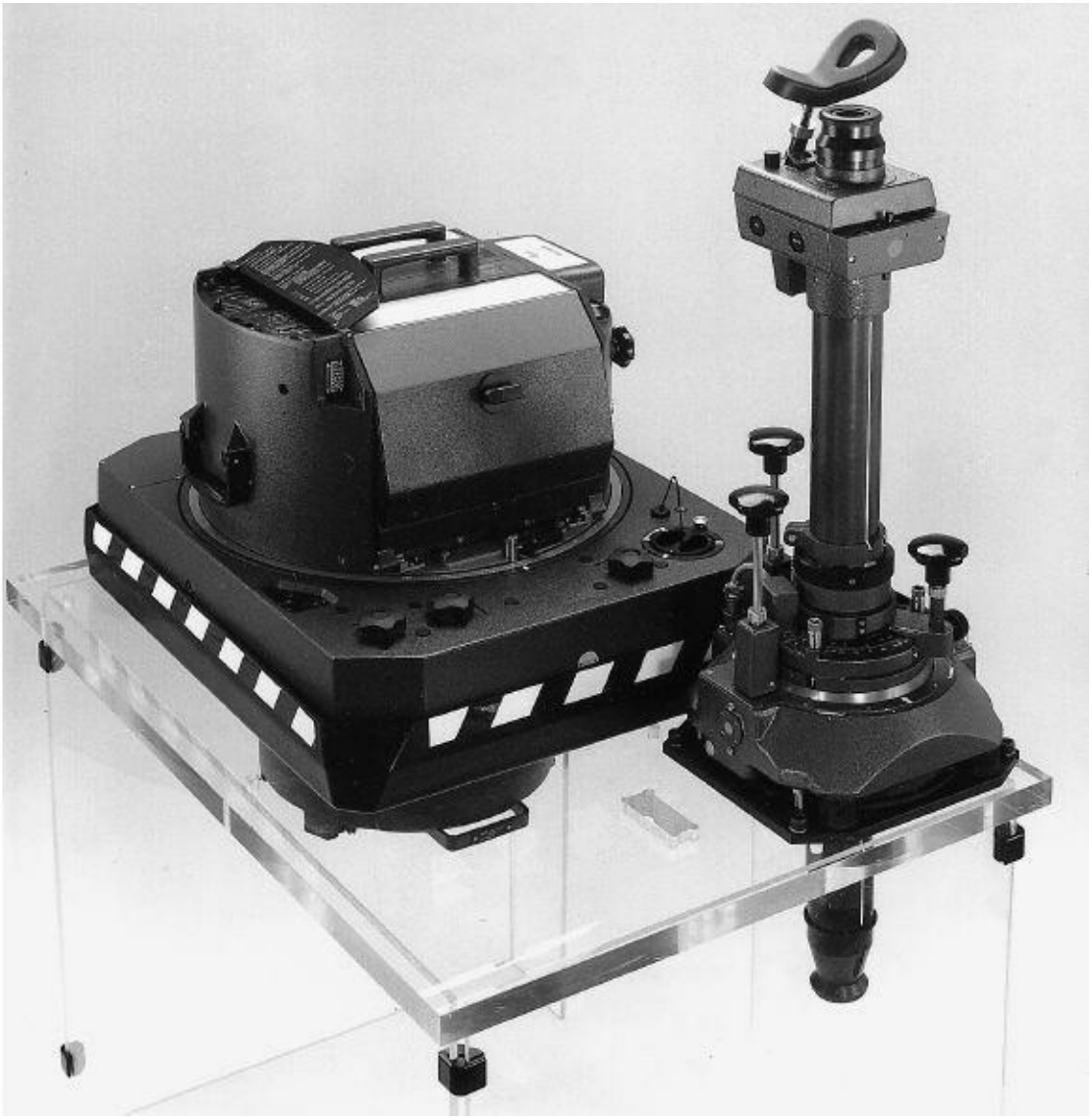


FIGURE 3-2 Leica RC30 aerial mapping camera. (*Courtesy LH Systems, LLC.*)

The Hasselblad camera of Fig. 3-3 is a small-format single-lens frame camera which has been employed extensively for space photography. It uses 70-mm film and can be obtained with various focal-length lenses.

FIGURE 3-3 Hasselblad 500 ELX camera. (*Courtesy University of Florida.*)

3-3 Main Parts of Frame Aerial Cameras

Although all frame aerial cameras are somewhat different in construction, they are enough alike that a general description can be given which adequately encompasses all of them. The three basic components or assemblies of a frame aerial camera, as shown in the generalized cross section of Fig. 3-4, are the *magazine*, the *camera body*, and the *lens cone assembly*.

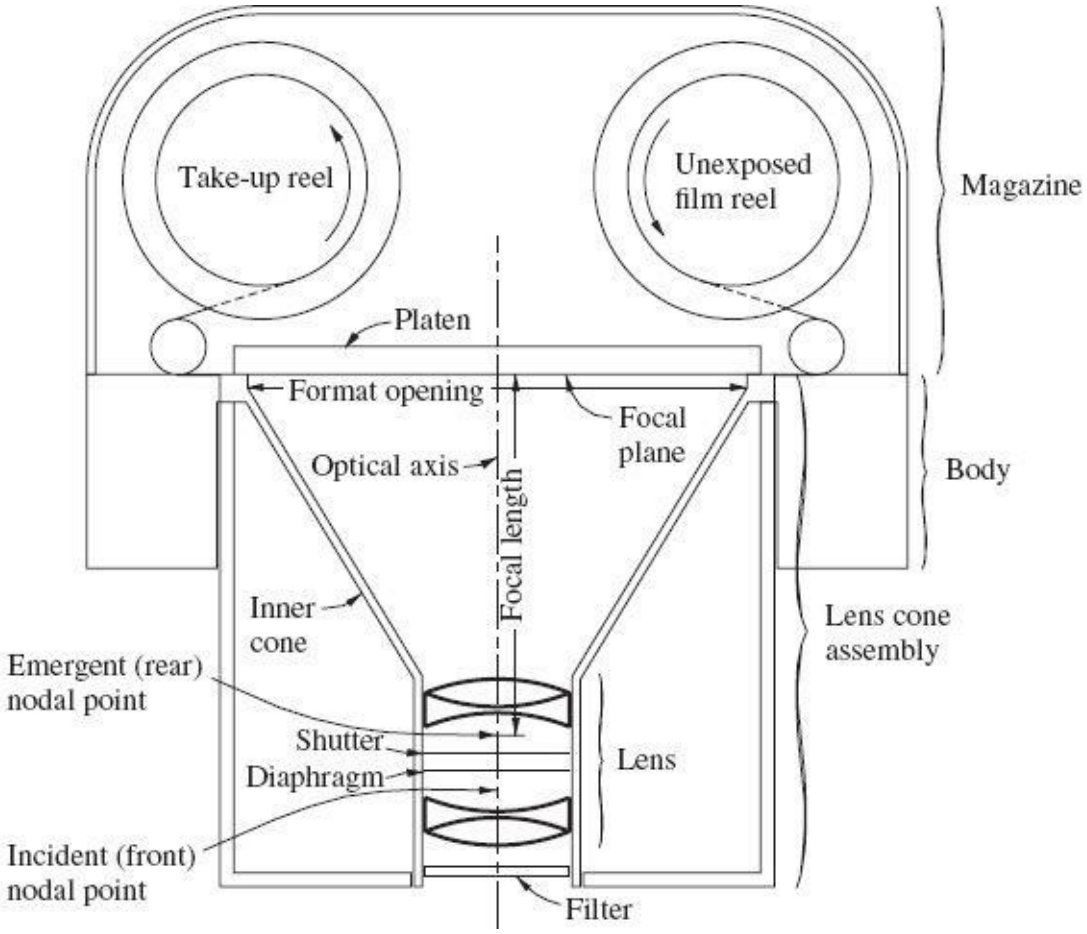


FIGURE 3-4 Generalized cross section of a frame aerial camera.

3-3-1 Camera Magazine

The camera magazine houses the reels which hold exposed and unexposed film, and it also contains the *film-advancing* and *film-flattening* mechanisms. Film flattening is very important in aerial cameras, for if the film buckled during exposure, the image positions on the resulting photographs would be incorrect. Film flattening may be accomplished in any of the following four ways: (1) by applying tension to the film during exposure; (2) by pressing the film firmly against a flat focal-plane glass that lies in front of the film; (3) by applying air pressure into the airtight camera cone, thereby forcing the film against a flat plate lying behind the focal plane; or (4) by drawing the film tightly up against a vacuum plate or platen whose surface lies in the focal plane. The vacuum system has proved most satisfactory and is the most widely used method of film flattening in aerial cameras. A focal-plane glass in front of the film is objectionable because image positions are distorted due to refraction

of light rays passing through the glass (see Sec. 2-2). These distortions can be determined through calibration, however, and their effect eliminated in subsequent photogrammetric operations.

3-3-2 Camera Body

The camera body is a one-piece casting which usually houses the drive mechanism. The drive mechanism operates the camera through its cycle; the cycle consists of (1) advancing the film, (2) flattening the film, (3) cocking the shutter, and (4) tripping the shutter. Power for the drive mechanism is most commonly provided by an electric motor. The camera body also contains carrying handles, mounting brackets, and electrical connections.

3-3-3 Lens Cone Assembly

The lens cone assembly contains a number of parts and serves several functions. Contained within this assembly are the *lens, shutter,* and *diaphragm* (see Fig. 3-4). With most mapping cameras, the lens cone assembly also contains an *inner cone* or *spider*. The inner cone rigidly supports the lens assembly and focal plane in a fixed relative position. This fixes the so-called elements of *interior orientation* of the camera. These elements are carefully determined through camera calibration (see Sec. 3-10) so that they are available for photogrammetric calculations. The inner cone is made of metal having a low coefficient of thermal expansion, so that changes in operating temperatures do not upset the calibration. In some aerial cameras which do not have inner cones, the body and outer lens cone act together to hold the lens with respect to the focal plane.

The *camera lens* is the most important (and most expensive) part of an aerial camera. It gathers light rays from the object space and brings them to focus in the focal plane behind the lens. Lenses used in aerial cameras are highly corrected compound lenses consisting of several elements. The lens of Fig. 2-9, for example, is the 88-mm-focal-length *SAGA-F* used in the superwide-angle lens cone of the Leica RC30 aerial mapping camera.

The *filter* serves three purposes: (1) It reduces the effect of atmospheric haze, (2) it helps provide uniform light distribution over the entire format, and (3) it protects the lens from damage and dust.

The *shutter* and *diaphragm* together regulate the amount of light which will expose the photograph. The shutter controls the length of time that light is permitted to pass through the lens. Shutters are discussed in detail in Sec. 3-5. As discussed in Sec. 2-5, the diaphragm regulates the *f*-stops of the camera by varying the size of the aperture to control the amount of light passing through the lens. Typically *f*-stops of aerial cameras range from about *f*-4 down to *f*-22. Thus, for a nominal 152-mm-focal-length lens, the diameter of the aperture ranges from about 38 mm at *f*-4 to about 7 mm at *f*-22. The diaphragm is normally located in the airspace between the lens elements of an aerial camera and consists of a series of leaves which can be rotated to vary the size of the opening.

3-4 Focal Plane and Fiducial Marks

The *focal plane* of an aerial camera is the plane in which all incident light rays are brought to focus. In aerial photography, object distances are great with respect to image distances. Aerial cameras therefore have their focus fixed for infinite object distances. This is done by setting the focal plane as exactly as possible at a distance equal to the focal length behind the rear nodal point of the camera lens. The focal plane is defined by the upper surface of the focal-plane frame. This is the surface upon which the film emulsion rests when an exposure is made.

Camera *fiducial marks* are usually four or eight in number, and they are situated in the middle of the sides of the focal plane opening, in its corners, or in both locations. Blinking lamps cause these

marks to be exposed onto the negative when the picture is taken. Modern mapping cameras generally expose the fiducials at the midpoint of the duration that the shutter is open. This defines the *instant of exposure*, which is critical when incorporating airborne GPS control. This topic is discussed in Chap. 17. The aerial photographs of Figs. 1-2, 1-6, and 1-7 have four corner and four side fiducial marks.

Fiducial marks (or fiducials) serve to establish a reference *xy* photo coordinate system for image locations on the photograph. In essence, fiducials are two-dimensional control points whose *xy* coordinates are precisely and accurately determined as a part of camera calibration (see Sec. 3-10). Lines joining opposite fiducials intersect at a point called the *indicated principal point*, and aerial cameras are carefully manufactured so that this occurs very close to the true *principal point*, which is defined as the point in the focal plane where a line from the rear nodal point of the camera lens, perpendicular to the focal plane, intersects the focal plane. As will be demonstrated in subsequent chapters, it is an exceedingly important reference point in photogrammetric work. Besides providing a coordinate reference for the principal point and image points, fiducials allow for correction of film distortion (shrinkage and expansion) since each photograph contains the images of these stable control points (see Sec. 4-9).

Cameras have been perfected to compensate for image motion that occurs because of the forward movement of the aircraft during the time that the shutter is open. *Forward-motion compensation* (FMC) is usually accomplished by moving the film slightly across the focal plane during exposure, in the direction of, and at a rate just equal to, the rate of image movement. In cameras equipped with FMC, since the film moves with respect to the focal plane, it is important for fiducial lamps to blink for only an instant so that the images of the fiducial marks are sharp and distinct.

Example 3-1

An aerial camera with forward-motion compensation and a 152.4-mm focal length is carried in an airplane traveling at 200 kilometers per hour (km/h). If the flying height above the terrain is 3500 m and if the exposure time is $\frac{1}{500}$ s, what distance (in millimeters) must the film be moved across the focal plane during exposure in order to obtain a clear image?

Solution The distance *D* traveled by the airplane during exposure is

$$D = \left(200\frac{\text{km}}{\text{h}}\right)\left(\frac{1}{500}\text{s}\right)\left(\frac{1\text{ h}}{3600\text{ s}}\right)\left(\frac{1000\text{ m}}{1\text{ km}}\right) = 0.11\text{ m}$$

The distance *d* that the image moves during exposure [based on scale Eq. (6-1)] is

$$d = 0.11\text{ m}\left(\frac{152.4\text{ mm}}{3500\text{ m}}\right) = 0.005\text{ mm} \quad \blacktriangle$$

3-5 Shutters

As noted in the previous section, because the aircraft carrying a camera is typically moving at a rapid speed, images will move across the focal plane during exposure. If exposure times are long or flying heights low, blurred images may result. It is important, therefore, that the shutter be open for a very short duration when aerial photographs are taken. Short exposure times also reduce the detrimental effects of aircraft vibrations on image quality. The shutter speeds of aerial cameras typically range

from about $\frac{1}{100}$ to $\frac{1}{1000}$ s. Shutters are designed to operate efficiently so that they open instantaneously, remain open the required time, and then instantaneously close, thus enabling the most uniform exposure possible over the format. (Other effects such as vignetting and lens falloff, as discussed in Sec. 2-3, cause nonuniform exposure, which is unavoidable.)

There are a number of different types of camera shutters. Those used in aerial cameras are generally classified as either *between-the-lens* shutters or *focal-plane* shutters. Between-the-lens shutters are most commonly used in mapping cameras. These shutters are placed in the airspace between the elements of the camera lens, as illustrated in Fig. 3-4. Common types of between-the-lens shutters are the *leaf* type, *blade* type, and *rotating-disk* type. A schematic diagram of the leaf type is shown in Fig. 3-5. It consists usually of five or more leaves mounted on pivots and spaced around the periphery of the diaphragm. When the shutter is tripped, the leaves rotate about their pivots to the open position of Fig. 3-5*b*, remain open the desired time, and then snap back to the closed position of Fig. 3-5*a*. Some camera shutters use two sets of leaves, one for opening and the other for closing. This increases shutter efficiency, shutter speed, and shutter life.

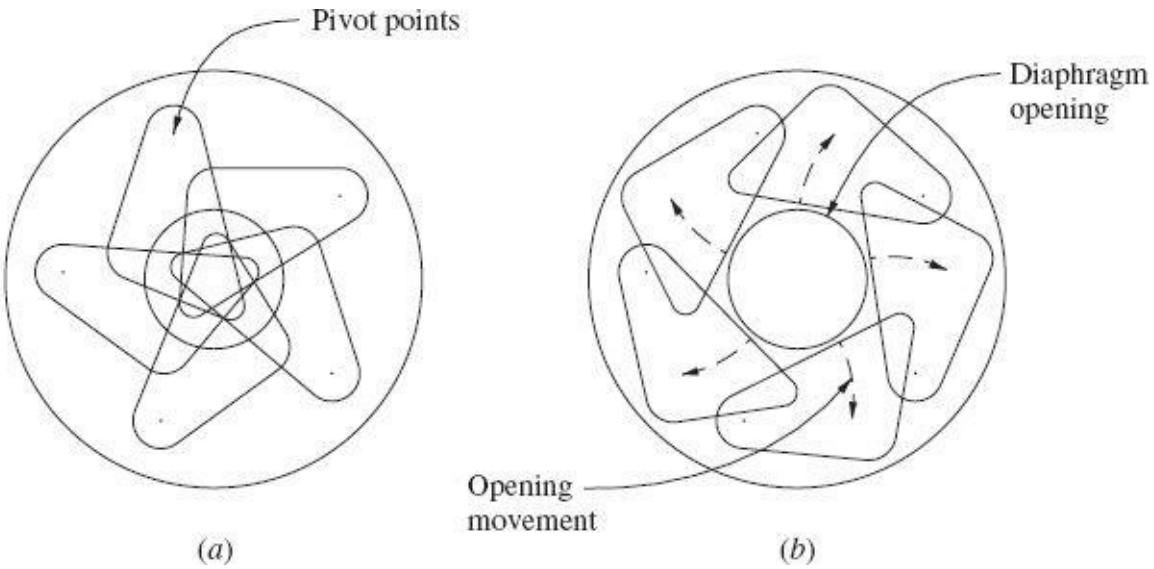


FIGURE 3-5 Schematic diagrams of a leaf-type shutter. (*a*) Shutter closed; (*b*) shutter open.

The blade-type shutter consists of four blades, two for opening and two for closing. Its operation is similar to that of a guillotine. When the shutter is triggered, the two thin *opening* plates or blades move across the diaphragm to open the shutter. When the desired exposure time has elapsed, two *closing* blades close it.

The rotating-disk type of shutter consists of a series of continuously rotating disks. Each disk has a cutaway section, and when these cutaways mesh, the exposure is made. The speed of rotation of the disks can be varied so that the desired exposure times are obtained. This type of shutter is very efficient because no starting or stopping of parts is required, as is with other types.

3-6 Camera Mounts

The camera mount is the mechanism used to attach the camera to the aircraft. Its purpose is to constrain the angular alignment of the camera so that the optical axis is vertical and the format is squarely aligned with the direction of travel. A minimal mount is equipped with dampener devices which prevent (or at least reduce) aircraft vibrations from being transmitted to the camera, and a

mechanism that allows rotation in azimuth to correct for *crab*. Crab is a disparity in the orientation of the camera in the aircraft with respect to the aircraft's actual travel direction. It is usually the result of side winds which cause the aircraft's direction of heading to deviate from its actual travel direction, as shown in Fig. 3-6*a*. Crab can be of variable amounts, depending on the wind velocity and direction. It has the undesirable effect of reducing the stereoscopic ground coverage (see Sec. 1-4) of aerial photos, as shown in Fig. 3-6*b*. Figure 3-6*c* shows the ground coverage when the camera has been rotated within the mount in the aircraft to make two sides of the format parallel to the actual direction of travel.

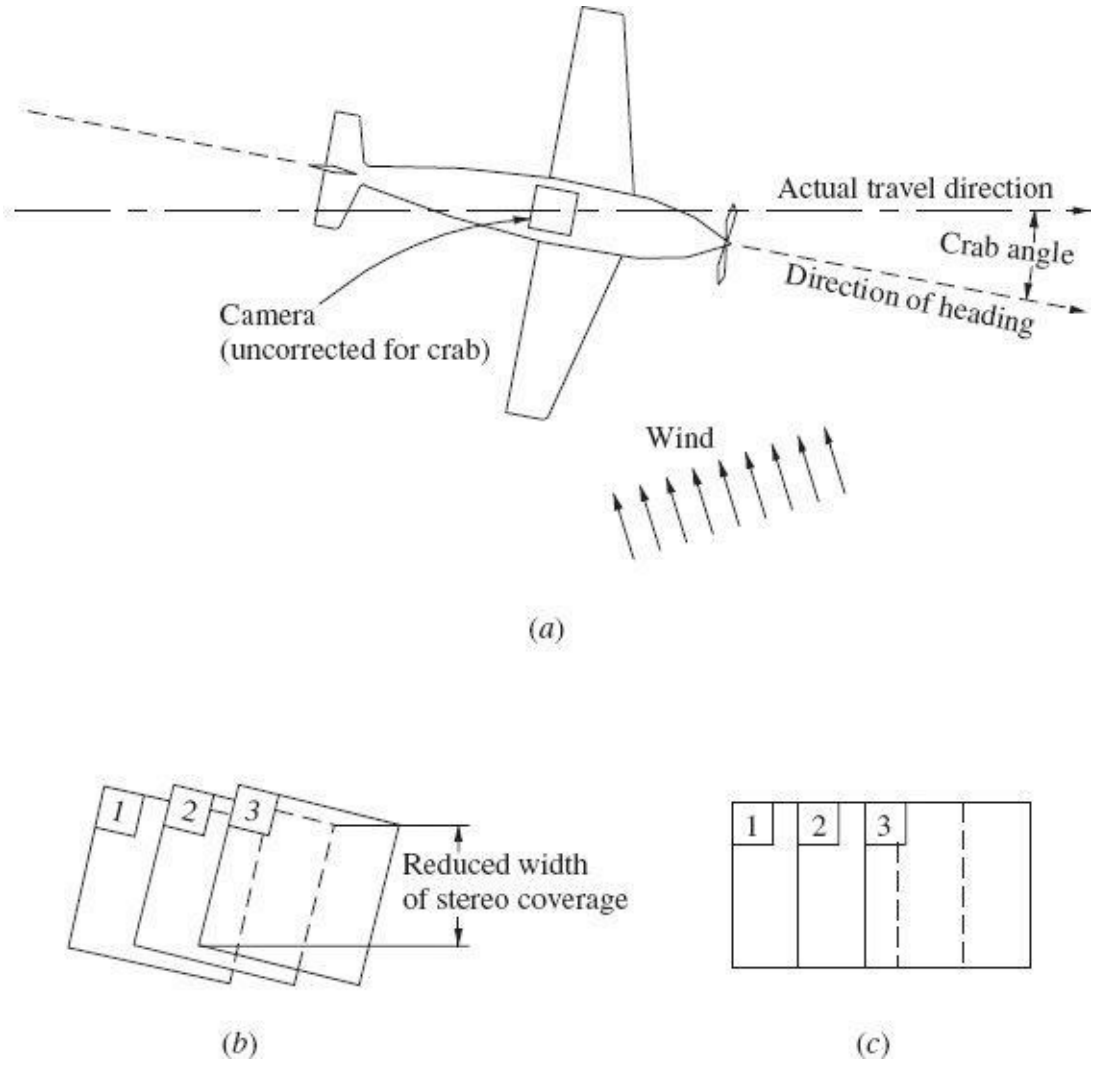


FIGURE 3-6 (*a*) Camera exposing aerial photography with crab present. (*b*) Crabbed overlapping aerial photographs. (*c*) Overlapping aerial photographs with no crab.

More elaborate mounts like the Leica PAV 80, shown in Fig. 3-7, provide gyro stabilization of the camera. Gyroscopic devices in the housing or in an external device sense the rotational movements of the aircraft, which in turn are counteracted by microprocessor-controlled motors that keep the camera properly oriented. Control is provided in three directions: rotation about the longitudinal axis (roll), rotation about the transverse axis (pitch), and rotation about the optical axis (yaw or drift). In addition to simply counteracting the aircraft's rotational movements, the three rotations are measured and can be recorded at the instant of exposure. These three rotational quantities are essential for proper data reduction when using airborne GPS control. When combined with forward-motion compensation, a gyro-stabilized mount results in the sharpest images by minimizing image movement

during exposure.

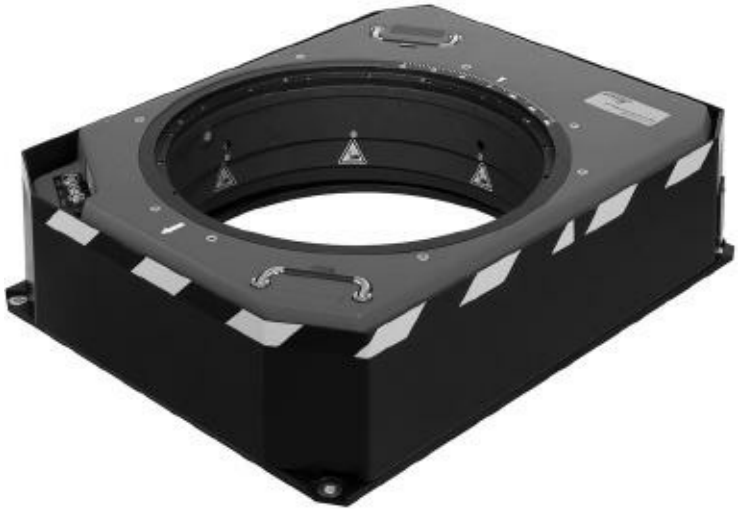


FIGURE 3-7 Leica PAV 80 gyro-stabilized aerial-camera mount. (*Courtesy Leica Geosystems*)

3-7 Camera Controls

Camera controls are those devices necessary for operating the camera and varying settings according to conditions at the time of photography. The *intervalometer* is a device which automatically trips the shutter and actuates the camera cycle as required by the flight plan. Early intervalometers could be set to automatically make exposures at fixed intervals of time. The time interval depended upon the camera focal length and format size, desired end lap (see Sec. 1-4), flying height above ground, and aircraft velocity. The disadvantage of this type of intervalometer is that with fixed time intervals, variations in end lap occur with variations in terrain elevation, flying height, or aircraft velocities. Subsequent intervalometer designs, include integrated automatic control unit that incorporates a GPS receiver, enabling the exposures to be made at preprogrammed locations as dictated by the flight plan.

Another aerial camera control device is the *exposure control*. This consists of an exposure meter which measures terrain brightness and correlates it with the optimum combination of aperture size and shutter speed, given a particular film speed and filter factor. Exposure control units are available which operate automatically, and they constantly vary camera settings to provide optimum exposures.

Modern camera controls consist of integrated navigation and sensor control units that are highly automated. By incorporating precise GPS and inertial navigation devices (see Sec. 16-6 to 16-8), these systems relieve the pilot and photographer of many routine in-flight monitoring and adjustment functions, resulting in reduced human error. In addition, flights can be made more efficient by reducing excessive side lap and avoiding gaps in ground coverage. The Flight and Sensor Control Management System shown in Fig. 3-8 can be used to automate a substantial portion of the image acquisition mission.

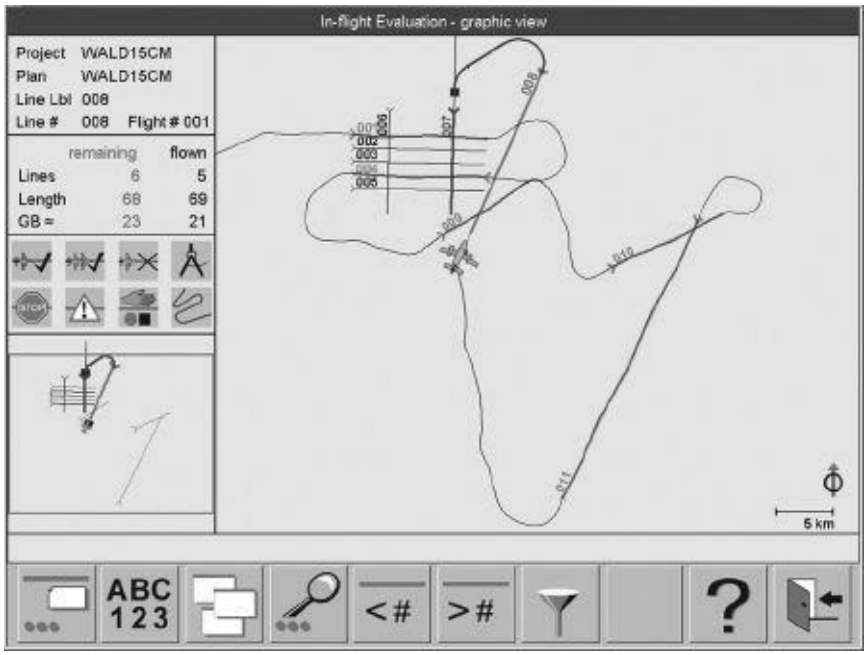


FIGURE 3-8 Flight and Sensor Control Management System (*Courtesy Leica Geosystems*)

3-8 Automatic Data Recording

Most modern aerial mapping cameras are equipped with a data-recording system which automatically produces pertinent data on the pictures. The usual data consist of date, flying altitude, calibrated focal length of the camera, photograph number, job identification, etc. This information is entered on a *data block*, which is exposed onto the film when the photograph is taken. The images of an automatic data-recording system are shown along the left border of the aerial photograph in Fig. 1-2. Automatic data recording is a convenience which saves time and prevents mistakes in later use of the photographs.

3-9 Digital Mapping Cameras

The majority of the material of this chapter to this point has pertained to traditional film cameras. As noted earlier, an alternative to film cameras is the digital imaging device. A digital image, as described in Sec. 2-12, is a rectangular array of pixels in which the brightness of a scene at each discrete location has been quantified. Rather than record the reflected electromagnetic energy through silver halide crystals as in a film emulsion, digital imaging devices use solid-state detectors to sense the energy. A common type of solid-state detector in current use is the *charge-coupled device* (CCD). Although there are a number of variations of CCD configurations, the basic operating principle is the same. At a specific pixel location, the CCD element is exposed to incident light energy, and it builds up an electric charge proportional to the intensity of the incident light. The electric charge is subsequently amplified and converted from analog to digital form. A large number of CCDs can be combined on a silicon chip in a one-dimensional or two-dimensional array. While other solid-state sensors such as Complimentary Metal-Oxide Semiconductor (CMOS) devices are used in some applications, the term CCD will be used throughout this text to represent all solid-state image sensors.

3-9-1 Digital-Frame Cameras

A digital-frame camera has similar geometric characteristics to a single-lens frame camera that employs film as its recording medium. It consists of a two-dimensional array of CCD elements, called a *full-frame sensor*. The sensor is mounted in the focal plane of a single-lens camera. Acquisition of

an image exposes all CCD elements simultaneously, thus producing the digital image. Figure 3-9 shows a full-frame sensor. It consists of an array of CCD elements approximately 14,600 × 17,200, and thus it produces an image having over 250 million pixels. Figure 3-10 shows a schematic illustration of a digital-frame camera capturing an image of the ground. Light rays from all points in the scene pass through the center of the lens before reaching the CCD elements, thus producing the same type of point-perspective image as would have occurred if film were used.

FIGURE 3-9 Solid-state CCD imaging array of 14,600 × 17,200 (250 million) pixels. (*Courtesy Teledyne DALSA.*) (See also color insert.)

Digital-frame cameras can be classified in terms of the number of pixels in the digital image. Currently the term *megapixel* (1 million pixels) is used for indicating common image sizes. Inexpensive digital cameras may have arrays of roughly 2500 rows and 3500 columns for a total of 2500 × 3500 = 8,750,000 pixels or 8.7 megapixels. Large arrays such as that of Fig. 3-9 may have 250 million or more. Current technology can produce chips with individual CCD elements approximately 5 μm in size or even smaller. The array of Fig. 3-9 has a 5.6 μm pixel size and thus can capture an 82 × 99 mm image in the focal plane of a camera.

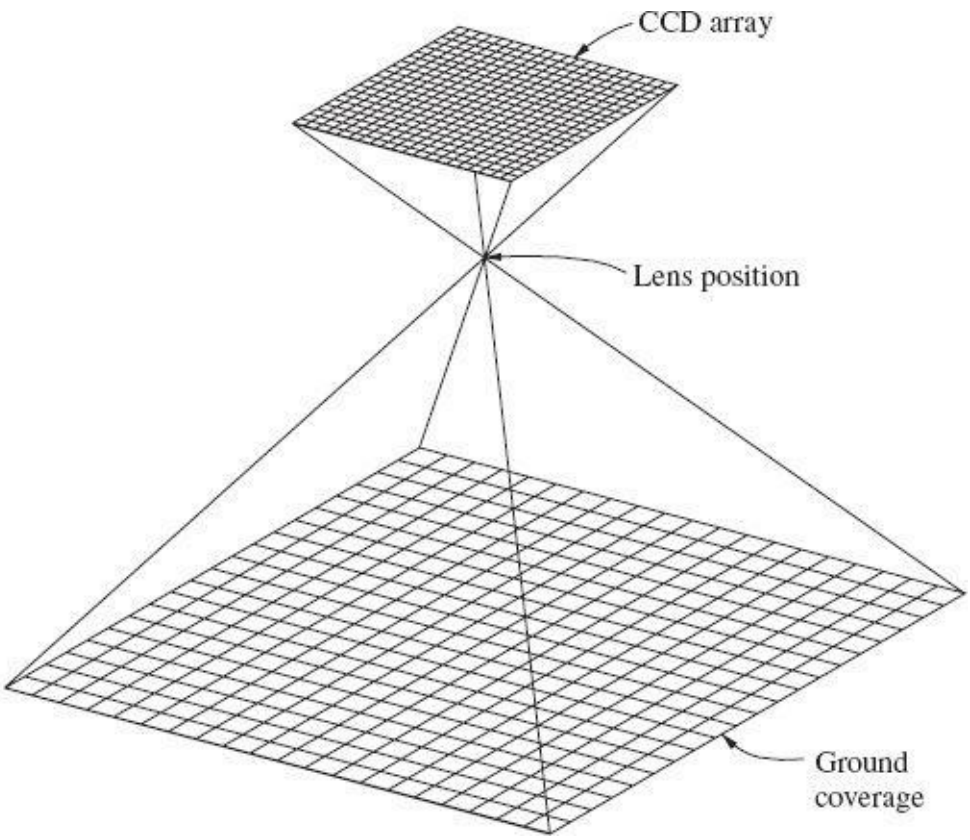


FIGURE 3-10 Geometry of a digital frame camera.

Figure 3-11 is a digital mapping camera that incorporates the CCD array shown in Fig. 3-9 along with a 112-mm-focal length lens. In addition to the ability to capture 250 megapixel panchromatic images, it is equipped to take lower resolution color and infrared images simultaneously.

FIGURE 3-11 Z/I DMC II-250 digital mapping camera. (*Courtesy Z/I Imaging*)

Some digital-frame cameras use an approach of combining multiple image subsets to produce a

composite image that has the geometric characteristics of a single-frame image. The camera in Fig. 1-3 employs this multiple image technique. An image is acquired by rapidly triggering four in-line sensors that combine to acquire nine image patches as shown in Fig. 3-12. The four sensors are aligned with the flight line and by precisely timing the image acquisitions to correspond to the aircraft velocity, all of the images will effectively be taken from a common location.

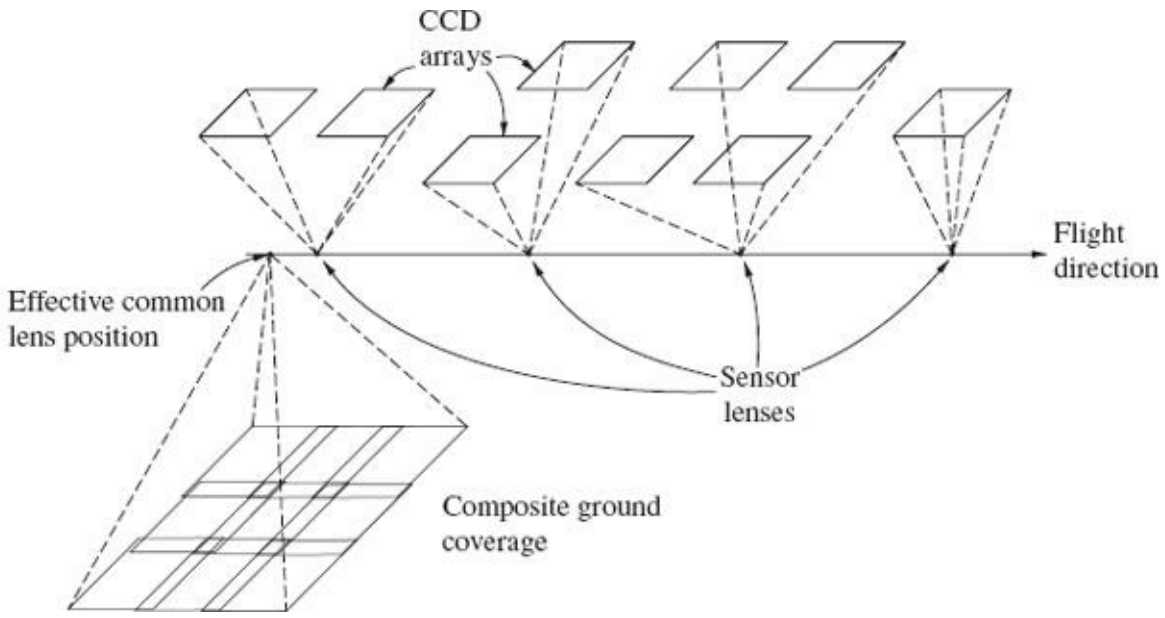


FIGURE 3-12 Multiple frame capture method of the UltraCam Eagle.

3-9-2 Linear Array Sensors

The geometric characteristics of a linear array sensor are different from those of a single-lens frame camera. At first glance, an image obtained from a linear array sensor may appear to be the same as a digital frame camera image, but there are subtle geometric differences. A linear array sensor acquires an image by sweeping a line of detectors across the terrain and building up the image. A linear array sensor consists of a one-dimensional array or strip of CCD elements mounted in the focal plane of a single-lens camera. Since the two-dimensional image is acquired in a sweeping fashion, the image is not exposed simultaneously. Figure 3-13 shows a schematic illustration of a linear array sensor capturing an image of the ground, assuming a smooth trajectory. At a particular instant, light rays from all points along a perpendicular to the vehicle trajectory pass through the center of the lens before reaching the CCD elements, thus producing a single row of the two-dimensional image. An instant later, the vehicle has advanced to its position for the next contiguous row, and the pixels of this row are imaged. The sensor proceeds in this fashion until the entire image is acquired. Since the image was acquired at a multitude of points along the line of the vehicle's trajectory, the resulting geometry is called *line perspective*.

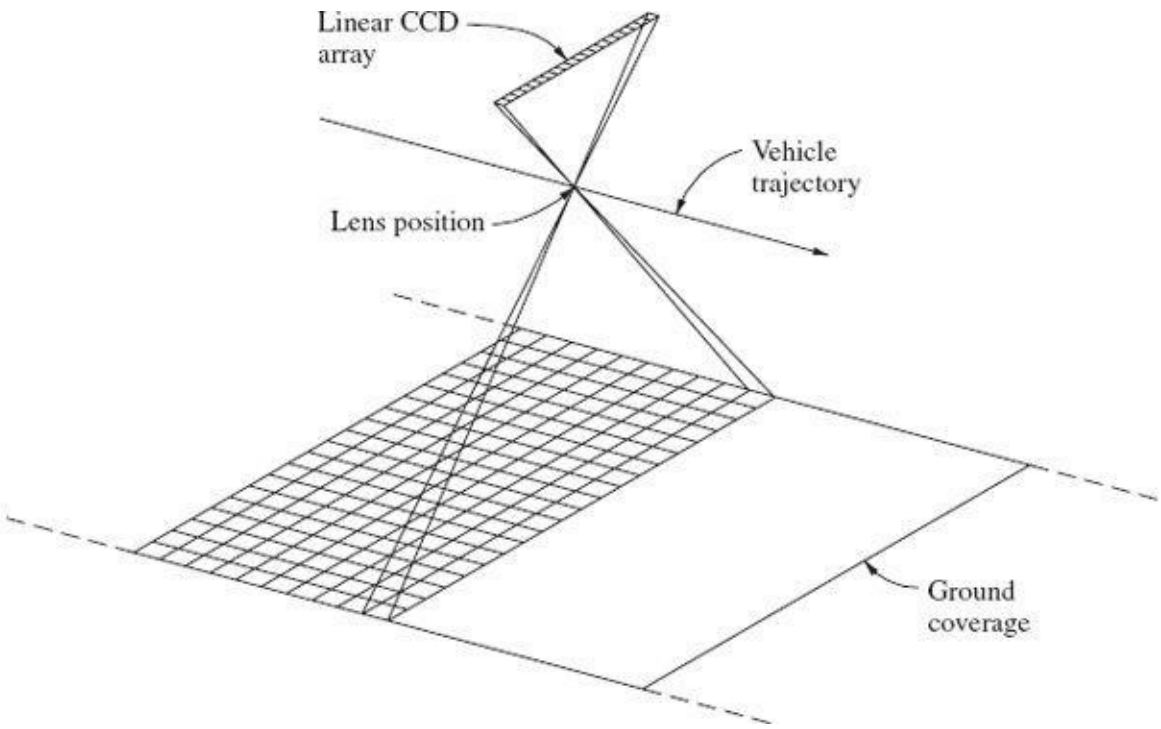


FIGURE 3-13 Geometry of a linear array sensor.

Acquisition of an image as depicted in Fig. 3-13 is characteristic of a linear array image captured from a satellite where the lack of atmospheric turbulence allows a smooth trajectory. When linear array images are acquired from a sensor carried by an airborne platform, atmospheric turbulence will cause the resulting image to be distorted due to pitching and rolling of the aircraft. In order to correct for this motion, simultaneous data must be collected by a GPS/INS system (See Sec. 16-6 to 16-8) to measure the position and angular attitude of the sensor. By processing the image to take into account the sensor motion, a nondistorted image can be produced. Figure 3-14(*a*) shows a raw image collected by a linear array camera similar to that of Fig. 1-4. Notice the wavy appearance of streets and other linear features in this uncorrected image. After rectifying the image based on the GPS/INS data, a corrected image shown in Fig. 3-14(*b*) is the result. This rectified image no longer shows the effects of the sensor motion and the linear features appear straight.

(*a*)

(*b*)

FIGURE 3-14 Raw image (*a*) from an airborne linear array sensor exhibiting distortion caused by air turbulence. Rectified image (*b*) obtained by correcting the raw image using GPS/INS measurements.

Another aspect of linear array sensors is that unlike frame sensors, they lack the ability to provide stereoscopic coverage in the direction of flight. This problem can be solved by equipping the sensor with multiple linear arrays pointed both forward and aft. The sensor of Fig. 1-4 has linear arrays that point forward, down, and backward as illustrated in Fig. 3-15. By using the images acquired by the forward-looking and backward-looking arrays, objects on the ground are imaged from two different vantage points, thus providing the required stereo view.

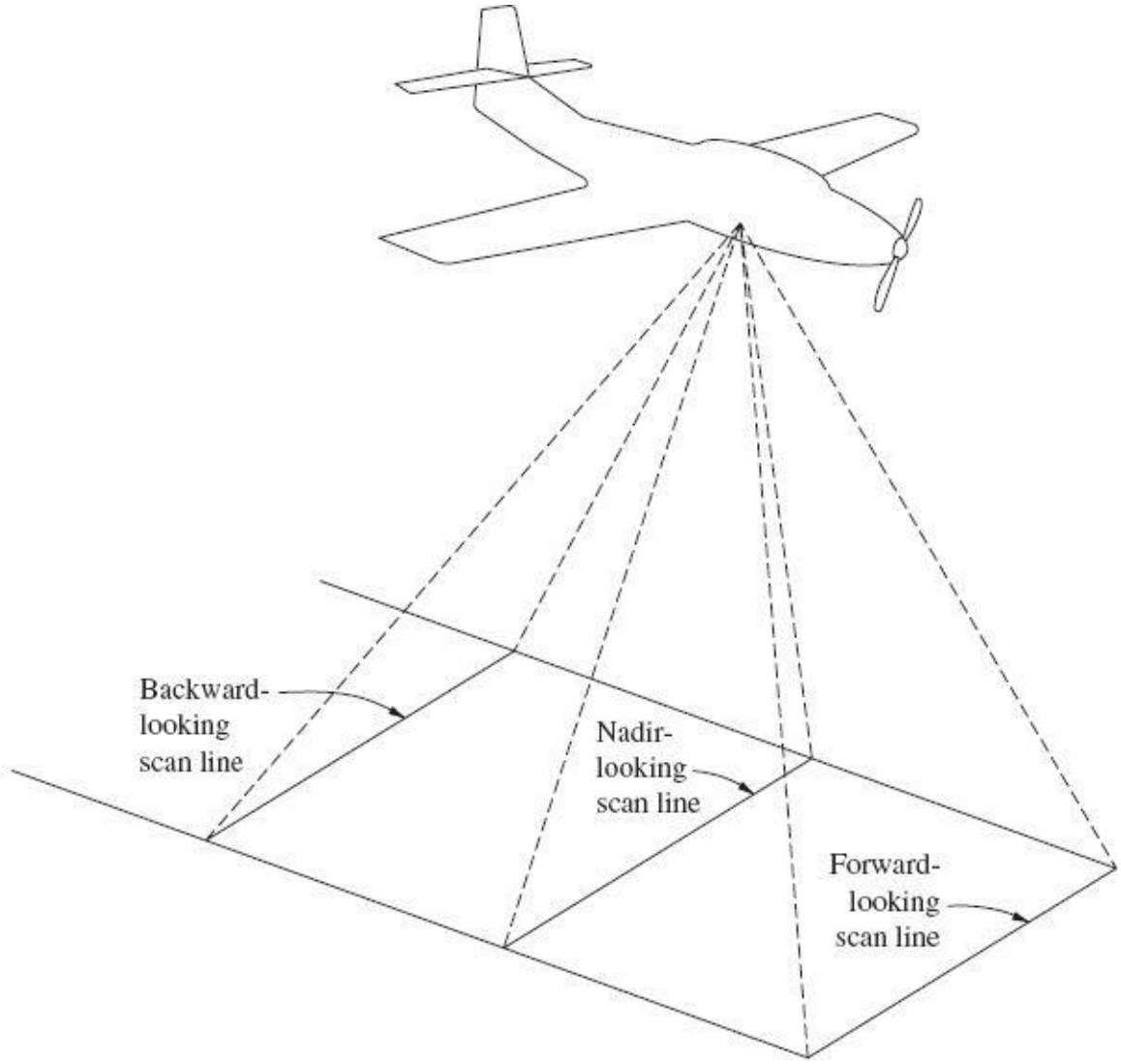


FIGURE 3-15 Geometry of a three-line scanner.

3-10 Camera Calibration

After manufacture and prior to use, aerial cameras are carefully calibrated to determine precise and accurate values for a number of constants. These constants, generally referred to as the *elements of interior orientation*, are needed so that accurate spatial information can be determined from photographs.

In general, camera calibration methods may be classified into one of three basic categories: *laboratory* methods, *field* methods, and *stellar* methods. Of these, laboratory methods are most frequently utilized and are normally performed by either camera manufacturers or agencies of the federal government. In one particular method of laboratory calibration, which uses a *multicollimator*, as well as in the field and stellar procedures, the general approach consists of photographing an array of targets whose relative positions are accurately known. Elements of interior orientation are then determined by making precise measurements of the target images and comparing their actual image locations with the positions they should have occupied had the camera produced a perfect perspective view. In another laboratory method, which employs a *goniometer*, direct measurements are made of projections through the camera lens of precisely positioned grid points located in the camera focal plane. Comparisons are then made with what the true projections should have been.

The elements of interior orientation which can be determined through camera calibration are as follows:

1. *Calibrated focal length* (CFL). This is the focal length that produces an overall mean distribution of lens distortion. Actually this parameter would be better termed *calibrated principal distance* since it represents the distance from the rear nodal point of the lens to the principal point of the photograph. When aerial mapping cameras are manufactured, this distance is set to correspond to the optical focal length of the lens as nearly as possible, hence the more common, though somewhat misleading, term *calibrated focal length*.
2. *Symmetric radial lens distortion*. This is the symmetric component of distortion that occurs along radial lines from the principal point. Although the amount may be negligible, this type of distortion is theoretically always present even if the lens system is perfectly manufactured to design specifications. Figure 3-16*a* shows a typical symmetric radial lens distortion pattern with magnitudes of distortion greatly exaggerated. Notice that distortion occurs in a direction inward toward, or outward from, the center of the image.

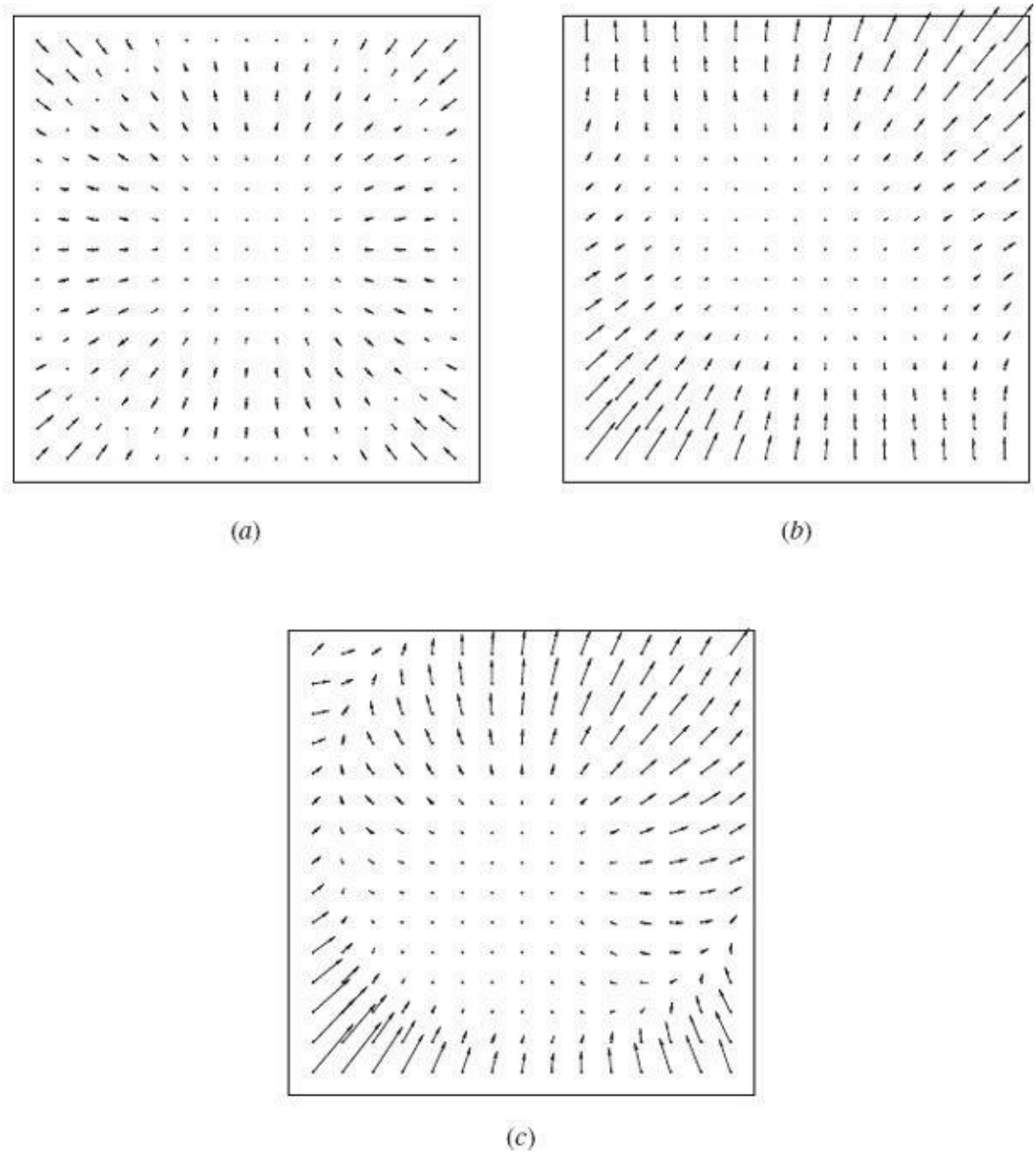


FIGURE 3-16 Lens distortion patterns: (*a*) symmetric radial, (*b*) decentering, and (*c*) combined symmetric radial and decentering.

3. *Decentering lens distortion*. This is the lens distortion that remains after compensation for symmetric radial lens distortion. Decentering distortion can be further broken down into *asymmetric radial* and *tangential* lens distortion components. These distortions are caused by imperfections in the manufacture and alignment of the lens system. Figure 3-16*b* shows a typical decentering distortion pattern, again with the magnitudes greatly exaggerated. Figure 3-16*c* shows a typical pattern of combined symmetric radial and decentering distortion.
4. *Principal point location*. This is specified by coordinates of the principal point given with respect to the *x* and *y* coordinates of the fiducial marks. (Although it is the intent in camera manufacture to place the fiducial marks so that lines between opposite pairs intersect at the principal point, there is always some small deviation from this ideal condition.) For a digital camera, the principal point is nominally located at the center of the CCD array, but calibration can determine the offset from this location.
5. *Fiducial mark coordinates*. These are the *x* and *y* coordinates of the fiducial marks which provide the two-dimensional positional reference for the principal point as well as images on the photograph. A digital camera does not have fiducial marks so these values are not determined from its calibration. Instead, the dimensions and effective shape of the CCD array are sometimes determined as part of the calibration. While the physical locations of the CCD elements tend to be highly accurate, the method by which the rows or columns of CCD elements are electronically sampled may cause a difference in the effective pixel dimensions in the *x* versus *y* directions.

In addition to the determination of the above elements of interior orientation, several other characteristics of the camera are often measured. *Resolution* (the sharpness or crispness with which a camera can produce an image) is determined for various distances from the principal point. Due to lens characteristics, highest resolution is achieved near the center, and lowest is at the corners of the photograph. *Focal-plane flatness* (deviation of the platen from a true plane) is measured by a special gauge. Since photogrammetric relationships assume a flat image, the platen should be nearly a true plane, generally not deviating by more than 0.01 mm. For digital cameras, direct measurement of the out-of-plane deviations is generally not feasible and therefore this distortion goes largely uncorrected. Often the *shutter efficiency*—the ability of the shutter to open instantaneously, remain open for the specified exposure duration, and close instantaneously—is also quantified.

3-11 Laboratory Methods of Camera Calibration

As noted in the previous section, the *multicollimator* method and the *goniometer* method are two types of laboratory procedures of camera calibration. The multicollimator method consists of photographing, onto a glass plate, images projected through a number of individual collimators mounted in a precisely measured angular array. A single collimator consists of a lens with a cross mounted in its plane of infinite focus. Therefore, light rays carrying the image of the cross are projected through the collimator lens and emerge parallel. When these light rays are directed toward the lens of an aerial camera, the cross will be perfectly imaged on the camera's focal plane because aerial cameras are focused for parallel light rays (those having infinite object distances).

A multicollimator for camera calibration consists of several individual collimators mounted in two perpendicular vertical planes (often, more than two planes are used). One plane of collimators is illustrated in Fig. 3-17. The individual collimators are rigidly mounted so that the optical axes of adjacent collimators intersect at known (measured) angles, such as θ_1 of Fig. 3-17. The *equivalent focal length* (EFL) is a computed value based on the distances from the center point *g* to each of the

four nearest collimator crosses (*h* and *f* are shown in Fig. 3-17 and the other two are in a perpendicular plane). Using the average of the four distances and the fixed angle θ_1 the EFL value is computed from the tangent ratio. The camera to be calibrated is placed so that its focal plane is perpendicular to the central collimator axis and the front nodal point of its lens is at the intersection of all collimator axes. In this orientation, image *g* of the central collimator, which is called the *principal point of autocollimation*, occurs very near the principal point, and also very near the intersection of lines joining opposite fiducials. The camera is further oriented so that when the calibration exposure is made, the collimator crosses will be imaged along the diagonals of the camera format, as shown in Fig. 3-18.

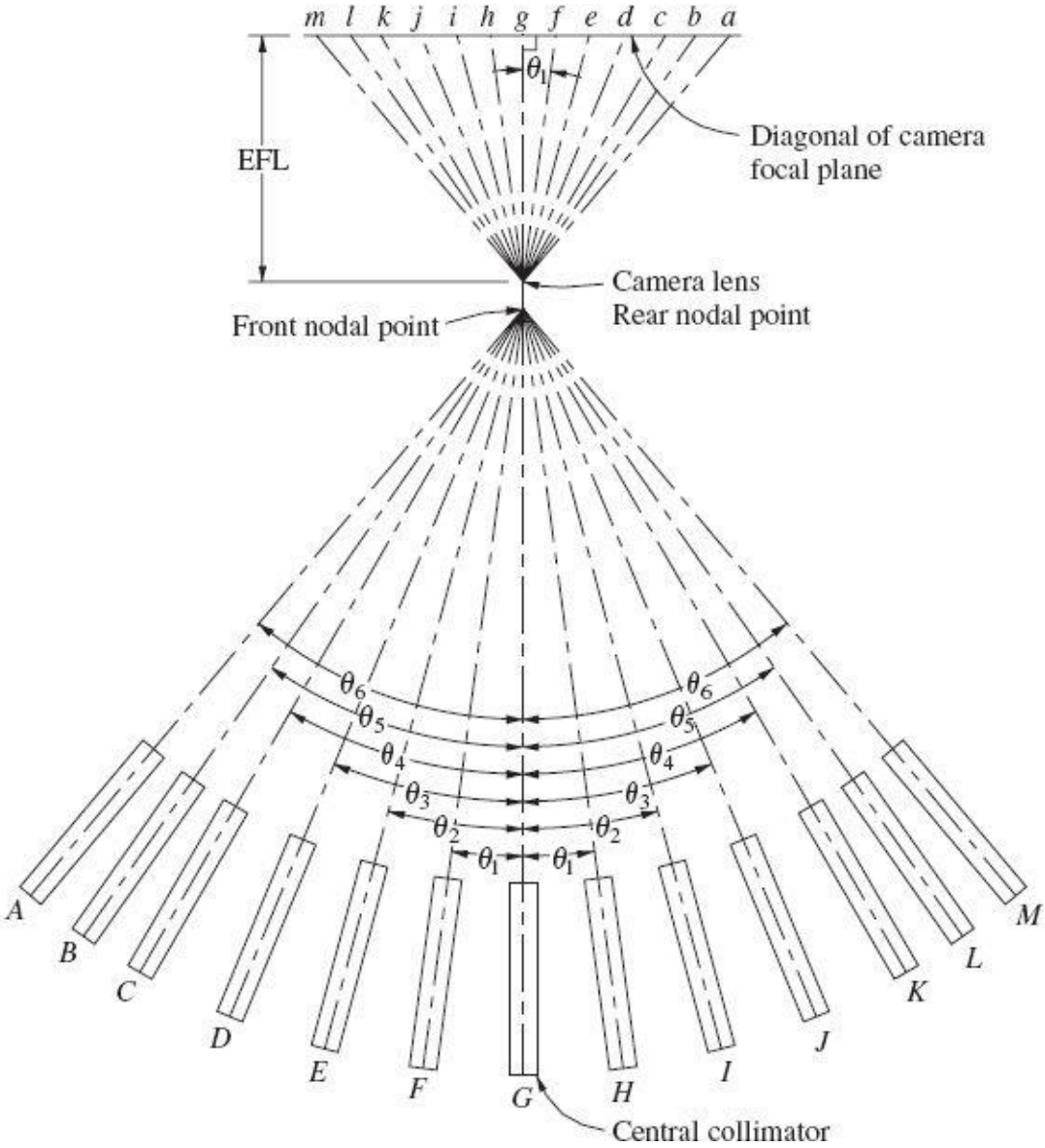


FIGURE 3-17 Bank of 13 collimators for camera calibration.

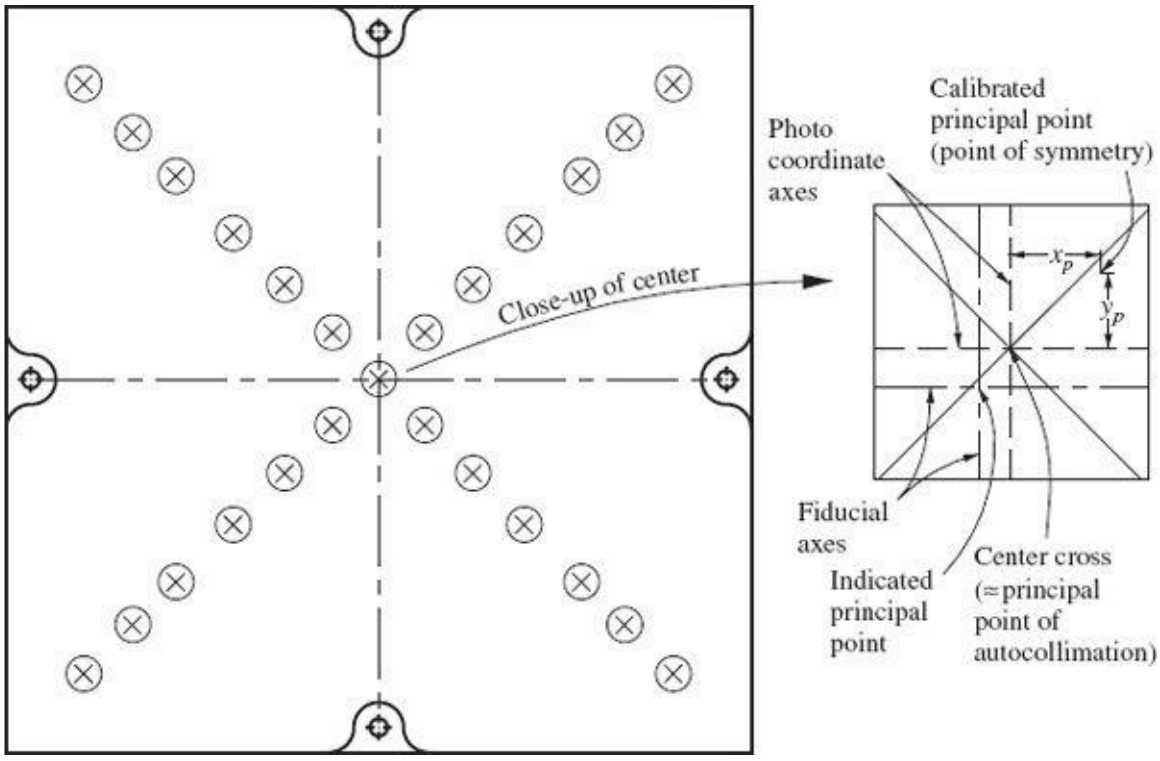


FIGURE 3-18 Images of photographed collimator targets and principal point definitions.

Figure 3-18 also contains a magnified view of the very center, which illustrates several key features. In the close-up, the fiducial lines are indicated which are simply lines joining opposite pairs of fiducials, and their intersection defines the *indicated principal point*. The position of the center collimator cross (principal point of autocollimation) typically serves as the origin of the photo coordinate system for film cameras. In a digital camera, the coordinates of the CCD elements are typically determined relative to an origin at one corner of the array. In this case, the principal point of autocollimation will have nonzero coordinates. The *calibrated principal point* (also known as the *point of best symmetry*) is the point whose position is determined as a result of the camera calibration. This point is the principal point that should be used to make the most precise and accurate photogrammetric calculations.

In determining the calibration parameters, a complex mathematical model is used which includes terms for the calibrated focal length and calibrated principal point coordinates as well as coefficients of symmetric radial lens distortion and decentering distortion. A least squares solution is performed which computes the most probable values for the above-mentioned terms and coefficients. The details of the calculations are beyond the scope of this text, but the fundamental concepts are presented in Chap. 19.

The goniometer laboratory procedure of camera calibration is very similar to the multicollimator method, but consists of centering a precision grid plate in the camera focal plane. The grid is illuminated from the rear and projected through the camera lens in the reverse direction. The angles at which the projected grid rays emerge are measured with a goniometer, a device similar to a surveyor's theodolite. CFL and lens distortion parameters are then computed with a mathematical model similar to that used in the multicollimator approach.

3-12 Stellar and Field Methods of Camera Calibration

Both the multicollimator and goniometer methods of laboratory camera calibration require expensive

and precise special equipment. An advantage of stellar and field methods is that this special equipment is not necessary. Several different stellar and field methods of camera calibration have been developed. In the stellar method, a target array consisting of identifiable stars is photographed, and the instant of exposure is recorded. Right ascensions and declinations of the stars can be obtained from an ephemeris for the precise instant of exposure so that the angles subtended by the stars at the camera station become known. Then these are compared to the angles obtained from precise measurements of the imaged stars. A drawback of this method is that since the rays of light from the stars pass through the atmosphere, compensation must be made for atmospheric refraction. On the other hand, there will be a large number of stars distributed throughout the camera format, enabling a more precise determination of lens distortion parameters.

Field procedures require that an array of targets be established and that their positions with respect to the camera station be measured precisely and accurately in three dimensions. This can be achieved conveniently using GPS methods. The targets are placed far enough from the camera station so that there is no noticeable image degradation. (Recall that an aerial camera is fixed for infinite focus.) In this configuration, the camera must be placed in a special apparatus such as a fixed tower, so that camera station coordinates are correctly related to target coordinates. This enables the CFL and principal point location to be determined as well as lens distortion parameters, even if the target configuration is essentially a two-dimensional plane. If the targets are well distributed in depth as well as laterally, accurate location of the camera is less important.

A variation of the field method described above, termed *inflight* camera calibration, can also be employed. In this approach, the aircraft carrying the camera makes multiple passes in different directions over a target range. Based on a high number of redundant measurements of target images, additional parameters (i.e., calibration parameters) can be calculated. This method has become more practical due to advancements in airborne GPS techniques which enable accurate camera station coordinates for each exposure (see Sec. 17-8). The in-flight method can also be generalized to the point where calibration parameters are determined in conjunction with the photographs taken during the actual job. This approach, known as *analytical self-calibration*, is further described in Sec. 19-4.

3-13 Calibration of Nonmetric Cameras

In certain situations where accuracy requirements and budgets are low, photogrammetrists may employ nonmetric cameras for acquisition of imagery. Nonmetric cameras are characterized by an adjustable principal distance, no film flattening or fiducial marks, and lenses with relatively large distortions. Calibration of a nonmetric camera allows at least some compensation to be made for these systematic errors.

The problems caused by the lack of fiducial marks and film flattening mechanisms presented a challenge when calibrating nonmetric film cameras. Fortunately nearly all nonmetric camera applications currently use digital cameras, so these issues are no longer present. However, the issue of the effective distorted shape of the CCD array must be addressed in order to obtain the highest degree of accuracy from digital cameras. As mentioned in Sec. 3-10, mathematical models can be used to correct for this effect, although the distortion from non-flatness remains.

A further complication in the calibration of nonmetric cameras arises when one is dealing with different focusing distances. Lens distortion values vary with different focus settings (i.e., different principal distances), thereby requiring a more general lens distortion model. This problem can be avoided by setting the focus of the camera to infinity during calibration as well as during normal use. That way, one set of lens distortion parameters can be determined to account for this consistent focus setting. If the camera is used for close-range work, the aperture size should be minimized in order to

produce the sharpest images. A word of caution regarding calibration of nonmetric cameras is warranted. The stability of focusable lenses—particularly zoom lenses—should be considered with suspicion. The calibration parameters for nonmetric cameras will often change significantly from day to day or even mission to mission, particularly if the camera has been bumped or subjected to vibrations. Nonmetric cameras should be calibrated both pre- and postmission to determine whether parameters have significantly changed. An even better approach would be to use analytical self-calibration as described in Sec. 19-4.

3-14 Calibrating the Resolution of a Camera

In addition to determining interior orientation elements, laboratory methods of camera calibration provide an evaluation of the camera's resolving power. As noted in Sec. 2-3, there are two common methods of specifying lens resolving power. One is a direct count of the maximum number of lines per millimeter that can be clearly reproduced by a lens; the other is the *modulation transfer function* (MTF) of the lens. The method of calibration employed to determine the line count consists of photographing resolution test patterns using a very high-resolution emulsion. The test patterns (an example is shown in Fig. 3-19) are comprised of numerous sets of *line pairs* (parallel black lines of varying thickness separated by white spaces of the same thickness). The measure of line thickness for each set is its number of line pairs per millimeter. Line thickness variations in a typical test pattern may range from 10 to 80 or more line pairs per millimeter. If the multicollimator method is used to calibrate a camera, the test patterns may be projected by the collimators simultaneously with the collimator crosses and imaged on the diagonals of the camera format. After the photograph is made, the resulting images are examined under magnification to determine the finest set of parallel lines that can be clearly resolved. The average of the four resolutions at each angular increment from the central collimator is reported in the calibration certificate. Another parameter generally reported is the *area-weighted average resolution* (AWAR), which is an indication of resolution over the entire format.

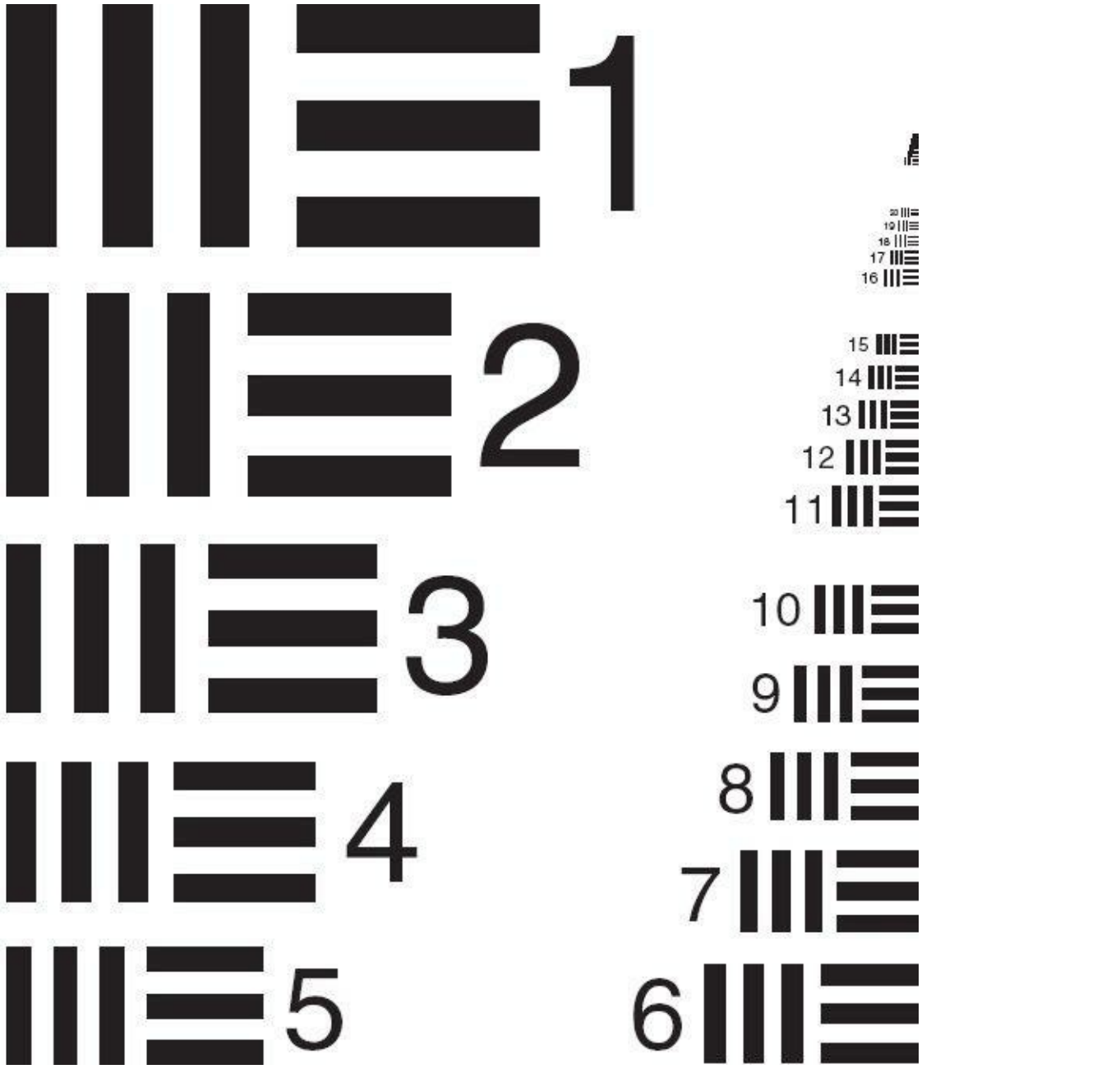


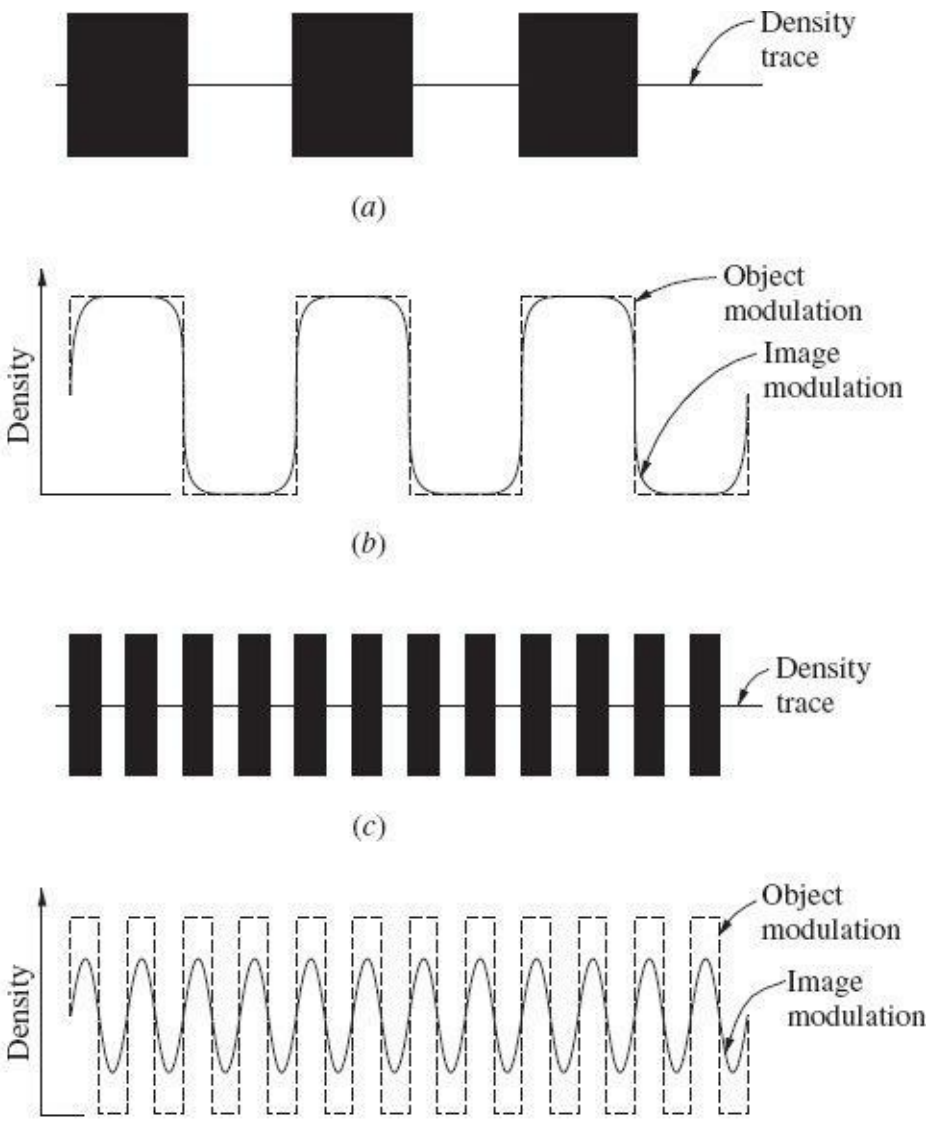
FIGURE 3-19 Resolution test pattern for camera calibration.

Whereas the above-described maximum-line-count method appears to be a relatively simple and effective way of quantifying resolving power, it is not without its shortcomings. In the line count procedure, with each succeedingly smaller test pattern, the sharpness of distinction between lines and spaces steadily diminishes, and the smallest pattern that can clearly be discerned becomes somewhat subjective. The preferred measure of resolution is the modulation transfer function.

A fundamental concept involved in quantifying the modulation transfer function is the notion of *spatial frequency*. Spatial frequency is a measure of the number of cycles of a sinusoidal wave per unit distance. An analogy can be drawn from audio signals where frequency concerns the number of sound waves per unit time. The units of audio frequency are generally specified in cycles per second, or hertz (abbreviated Hz), whereas units of spatial frequency are typically given in terms of cycles per millimeter. Spatial frequency is directly related to the count of line pairs per millimeter discussed above. A black-and-white line pair corresponds to the up-and-down pulse of a sine wave and thus can be defined as one cycle of a wave. Therefore the number of line pairs per millimeter is equivalent to cycles per millimeter, or spatial frequency. Images that contain areas of rapidly changing levels of

brightness and darkness have high spatial frequency, whereas images that contain areas of gently changing levels have low spatial frequency.

To determine modulation transfer, density scans using a photogrammetric scanner (see Sec. 4-6) are taken in a single trace across test patterns similar to those used in the line count procedure, as shown in Fig. 3-20*a* and *c*. For heavy lines with wide spacing, the actual distribution of density (brightness variations) across the object pattern would appear as the dashed lines shown in Fig. 3-20*b*, whereas brightness distributions measured with a densitometer across the image of this pattern would appear as the solid lines. Note that the edges of the image patterns are rounded somewhat in Fig. 3-20*b*, but the amplitude of brightness differences is the same as that for the original object. Thus at this spatial frequency of the pattern, modulation transfer is said to be 100 percent. Figure 3-20*c* shows an object pattern at a frequency four times that of the pattern shown in Fig. 3-20*a*. The density distributions of the object and resulting image of this higher-frequency pattern are shown in Fig. 3-20*d*. Note that in this figure, not only are the edges rounded, but also the amplitude of brightness differences is about one-half that of the original object. This indicates a modulation transfer of 50 percent from object to image. Actually, Fig. 3-20 is a somewhat simplified illustration of the quantification of modulation transfer. In the rigorous determination of modulation transfer, exposure values (rather than densities), which have a logarithmic relationship to density, are employed. The reader may consult references at the end of this chapter for more details on the modulation transfer function.



(*d*)

FIGURE 3-20 (*a*) Test object at low spatial frequency with density trace. (*b*) Density modulation of object (dashed) and image (solid). (*c*) Test object at high spatial frequency with density trace. (*d*) Density modulation of object (dashed) and image (solid). [Note that in part (*b*), the amplitude of the image modulation is the same as that of the object, corresponding to 100 percent modulation transfer. In (*d*) however, amplitude of the image modulation is one-half that of the object, corresponding to reduced modulation transfer.]

By measuring densities across many patterns of varying spatial frequencies, and plotting the resulting modulation transfer percentages on the ordinate versus corresponding spatial frequencies on the abscissa, a curve such as that illustrated in Fig. 3-21 is obtained. This curve is the modulation transfer function. The MTF has a number of advantages over the simple line count method. It is a very sensitive indicator of *edge effects*, and it also affords the capability of predicting the resolution that may be expected at any given degree of detail. Furthermore, MTF curves can be combined for different lenses, films, and film processes; thus, it is possible to estimate the combined effects of any given imaging system. For these reasons, the MTF has become the preferred method of expressing resolution.

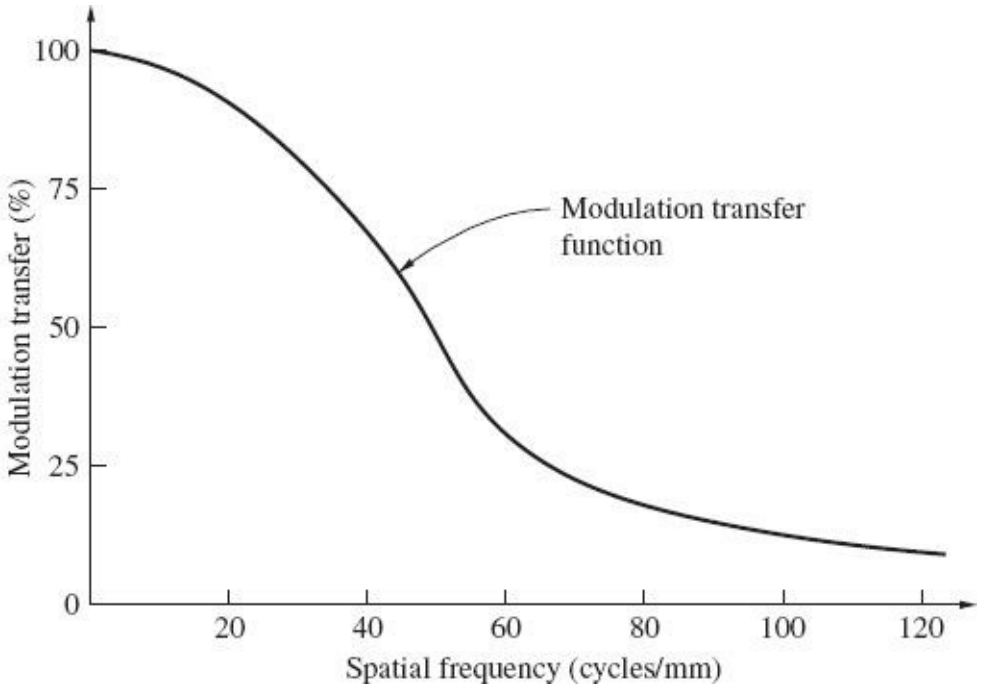


FIGURE 3-21 Curve of modulation transfer function (MTF).

The upper limit of resolution for a digital frame camera is absolutely fixed because of sampling into discrete elements. Since a full cycle of a wave in terms of spatial frequency must consist of a dark-to-light transition (line pair), two CCD elements are the minimum number that can capture information at the highest frequency. Thus the maximum spatial frequency (best resolution) at image scale that can be detected is

$$f_{\max} = \frac{1}{2w} \tag{3-2}$$

where w = width between centers of adjacent CCD elements

f_{max} = maximum detectable frequency

Example 3-2

A digital frame camera consists of a 5120 × 5120 array of CCD elements at a pixel size of 6 μm square. The nominal focal length of the camera is 40 mm. What is the maximum spatial frequency that can be detected (at image scale)? What is the angular field of view for this camera?

Solution By Eq. (3-2),

$$f_{\max} = \frac{1}{2w} = \frac{1}{2(0.006)} = 83 \text{ cycles/mm}$$

Format diagonal d can be calculated by

$$d = \sqrt{(5120 \times 0.006)^2 + (5120 \times 0.006)^2} = 43 \text{ mm}$$

By Eq. (3-1),

$$\alpha = 2\tan^{-1}\left(\frac{d}{2f}\right) = 2\tan^{-1}\left(\frac{43}{2 \times 40}\right) = 57^\circ \quad \blacktriangle$$

References

Abdel-Aziz, Y.: "Asymmetrical Lens Distortion," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 3, 1975, p. 337.

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004, chaps. 7 and 8.

———: *Manual of Remote Sensing*, 3d ed., Bethesda, MD, 1998.

Anderson, J. M., and C. Lee: "Analytical In-Flight Calibration," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 11, 1975, p. 1337.

Brock, G. C.: "The Possibilities for Higher Resolution in Air Survey Photography," *Photogrammetric Record*, vol. 8, no. 47, 1976, p. 589.

Brown, D. C.: "Close-Range Camera Calibration," *Photogrammetric Engineering*, vol. 37, no. 8, 1971, p. 855.

Carman, P. D.: "Camera Vibration Measurements," *Canadian Surveyor*, vol. 27, no. 3, 1973, p. 208.

Cramer, M., and N. Haala: "DGPF Project: Evaluation of Digital Photogrammetric Aerial-Based Imaging Systems—Overview and Results from the Pilot Center," *Photogrammetric Engineering and Remote Sensing*, vol. 76, no. 9, 2010, p. 1019.

Doyle, F. J.: "A Large Format Camera for Shuttle," *Photogrammetric Engineering and Remote Sensing*, vol. 45, no. 1, 1979, p. 737.

Fraser, C.S. and S. Al-Ajiouni: "Zoom-Dependant Camera Calibration in Digital Close-Range Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 72, no. 9, 2006, p.1017.

Fraser, C. S.: "Digital Camera Self-Calibration," *ISPRS Journal of Photogrammetry and Remote*

Sensing, vol. 52, 1997, p. 149.
Fraser, C. S., and M. R. Shortis: "Variation of Distortion within the Photographic Field," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 6, 1992, p. 851.
Habib, A. and M. Morgan: "Stability Analysis and Geometric Calibration for Off-the-Shelf Digital Cameras," Photogrammetric Engineering and Remote Sensing, vol. 71, no. 6, 2005, p. 733.
Honkavaara, E., J. Peltoniemi, E. Ahokas, R. Kuittinen, J. Hyyppä, J.Jaakkola, H. Kaartinen, L. Markelin, K. Nurminen, and J. Suomalainen: "A Permanent Test Field for Digital Photogrammetric Systems," *Photogrammetric Engineering and Remote Sensing*, vol. 74, no. 1, 2008, p. 95.
Hakkarainen, J.: "Image Evaluation of Reseau Cameras," *Photogrammetria*, vol. 33, no. 4, 1977, p. 115.
Karren, R. J.: "Camera Calibration by the Multicollimator Method," *Photogrammetric Engineering*, vol. 34, no. 7, 1968, p. 706.
Kenefick, J. F., M. S. Gyer, and B. F. Harp: "Analytical Self-Calibration," *Photogrammetric Engineering*, vol. 38, no. 11, 1972, p. 1117.
King, D., P. Walsh, and F. Ciuffreda: "Airborne Digital Frame Camera Imaging for Elevation Determination," *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 11, 1994, p. 1321.
Lei, F., and H. J. Tiziani: "A Comparison of Methods to Measure the Modulation Transfer Function of Aerial Survey Lens Systems from the Image Structures," *Photogrammetric Engineering and Remote Sensing*, vol. 54, no. 1, 1988, p. 41.
Light, D. L.: "The New Camera Calibration System at the U.S. Geological Survey," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 2, 1992, p. 185.
———: "Film Cameras or Digital Sensors? The Challenge ahead for Aerial Imaging," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 3, 1996, p. 285.
Markelin, L., E. Honkavaara, J. Peltoniemi, E. Ahokas, R. Kuittinen, J. Hyyppä, J. Suomalainen, and A. Kukko: "Radiometric Calibration and Characterization of Large Format Digital Photogrammetric Sensors in a Test Field," *Photogrammetric Engineering and Remote Sensing*, vol. 74, no. 12, 2008, p. 1487.
Merchant, D. C.: "Calibration of the Air Photo System," *Photogrammetric Engineering*, vol. 40, no. 5, 1974, p. 605.
Nielsen, V.: "More on Distortions by Focal Plane Shutters," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 2, 1975, p. 199.
Rampal, K. K.: "System Calibration of Metric Cameras," *ASCE Journal of the Surveying and Mapping Division*, vol. 41, no. SU1, 1978, p. 51.
Scholer, H.: "On Photogrammetric Distortion," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 6, 1975, p. 761.
Tayman, W. P.: "Calibration of Lenses and Cameras at the USGS," *Photogrammetric Engineering*, vol. 40, no. 11, 1974, p. 1331.
———: "User Guide for the USGS Aerial Camera Report of Calibration," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 5, 1984, p. 577.
Welch, R., and J. Halliday: "Imaging Characteristics of Photogrammetric Camera Systems," *Photogrammetria*, vol. 29, no. 1, 1973, p. 1.

Problems

3-1. List and briefly describe the three geometric categories of imaging devices.

3-2. List the requirements of a precision mapping camera.

3-3. What is the angular field of view of a camera having a 230-mm-square format and a 209-mm focal length?

3-4. Repeat Prob. 3-3 except that the format is 89-mm-square and the focal length is 112 mm.

3-5. For a camera having a 230-mm-square format, what range of focal lengths could it have to be classified as wide angle?

3-6. An aerial camera makes an exposure at a shutter speed of $\frac{1}{500}$ s. If aircraft speed is 390 km/h, how far does the aircraft travel during the exposure?

3-7. Repeat Prob. 3-6, except that the shutter speed is $\frac{1}{1000}$ s and aircraft speed is 170 mi/h.

3-8. An aerial camera with forward-motion compensation and a 152.4-mm focal length is carried in an airplane traveling at 280 km/h. If flying height above terrain is 3200 m and if the exposure time is $\frac{1}{250}$ s, what distance (in millimeters) must the film be moved across the focal plane during exposure in order to obtain a clear image?

3-9. Repeat Prob. 3-8, except that the focal length is 305 mm and the flying height above terrain is 5500 m.

3-10. Name and briefly describe the main parts of a frame aerial mapping camera.

3-11. Discuss briefly the different types of camera shutters.

3-12. What is the purpose of the camera mount?

3-13. What is the primary benefit of gyro-stabilized camera mounts?

3-14. What is crab, and how may it be caused?

3-15. Why is camera calibration important?

3-16. What are the elements of interior orientation that can be determined in camera calibration?

3-17. List and briefly describe the various definitions of *principal point*.

3-18. Briefly describe the advantages of using the modulation transfer function to quantify the resolution of a lens over the simple line pairs per millimeter threshold.

3-19. Illustrate and briefly describe the concept of spatial frequency.

3-20. A digital frame camera consists of a 4096 × 4096 array of CCD elements at a pixel size of 6.5 μm square. The nominal focal length of the camera is 40 mm. What is the maximum spatial frequency that can be detected (at image scale)? What is the angular field of view for this camera?

CHAPTER 4

Image Measurements and Refinements

4-1 Introduction

The solution of most photogrammetric problems generally requires some type of photographic measurement. For certain problems the measurements may simply be the lengths of lines between imaged points. However, rectangular coordinates of imaged points are the most common type of photographic measurement, and they are used directly in many photogrammetric equations. Photographic measurements are usually made on positives printed on paper, film, or glass, or in digital images manipulated on a computer. They could also be made directly on the negatives; however, this is seldom done because it can deface the imagery, and it is important to preserve the negatives for making additional prints. It is common, however, to make digital scans directly from the negatives, thus avoiding additional expense associated with making positive prints.

Equipment used for making photographic measurements varies from inexpensive, simple scales to very precise and complex machines that provide computer-compatible digital output. These various types of instruments and the manner in which they are used are described in this chapter. Because of several effects, there will be systematic errors associated with practically all photographic measurements. The sources of these errors and the manners by which they are eliminated are also discussed in this chapter.

4-2 Coordinate Systems for Image Measurements

For metric cameras with side fiducial marks, the commonly adopted reference system for photographic coordinates, is the rectangular axis system formed by joining opposite fiducial marks with straight lines, as shown in Fig. 4-1. The x axis is usually arbitrarily designated as the fiducial line most nearly parallel with the direction of flight, positive in the direction of flight. The positive y axis is 90° counterclockwise from positive x. The origin of the coordinate system is the intersection of fiducial lines. This point is often called the indicated principal point, as discussed in Sec. 3-4; for a precise mapping camera it is very near the true principal point. For the most precise work, since the lines joining opposite fiducials are not exactly perpendicular, the fiducial marks serve as control points from which the photo coordinate axis system can be determined.

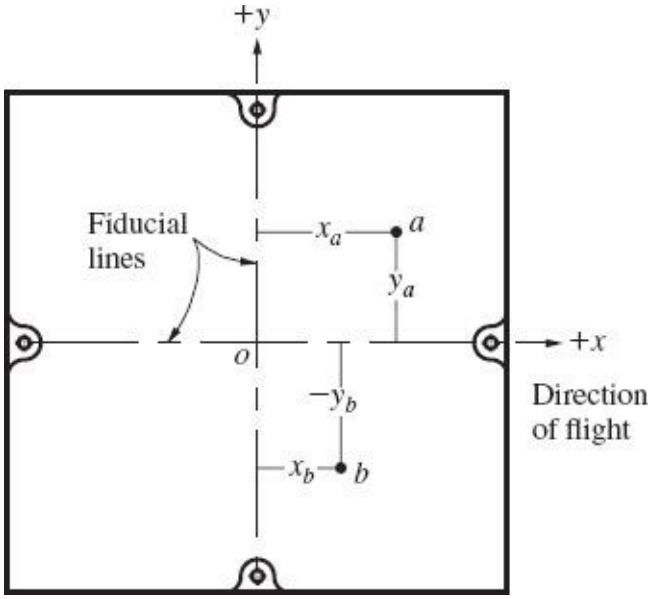


FIGURE 4-1 Photographic coordinate system based on side fiducials.

The position of any image on a photograph, such as point *a* of Fig. 4-1, is given by its rectangular coordinates x_a and y_a, where x_a is the perpendicular distance from the y axis to a and y_a is the perpendicular distance from the x axis to a. Similarly, the photographic position of image point b is given by its rectangular coordinates x_b and y_b.

It is very common for aerial cameras to have eight fiducials installed, in both side and corner locations. Figures 1-2, 1-6, and 1-7 show this fiducial mark configuration. The photographic coordinate system in this case is still defined as in Fig. 4-1. Eight fiducials enable somewhat more accurate corrections to be made for systematic errors in measured image coordinates.

On digital images, coordinates are often expressed as row and column numbers of individual pixels. If pixel dimensions are known, the dimensionless row and column values can be converted to linear measurements. Note that for many common image manipulation programs, pixel coordinates are given in the order of column (abscissa) followed by row (ordinate). This results in a "left-handed" coordinate system, which may lead to problems in certain calculations.

Rectangular coordinates are a very basic and useful type of photographic measurement. They are used in many different types of computations. As an example, they can be used to calculate the photo distances between points by using simple analytic geometry. Photographic distance *ab* of Fig. 4-1, for example, may be calculated from rectangular coordinates as follows:

$$ab = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \tag{4-1}$$

4-3 Simple Scales for Photographic Measurements

There are a variety of simple scales available for photographic measurements. The particular choice of measuring instrument will depend upon the accuracy required for the photogrammetric problem at hand. If a low order of accuracy is acceptable, an ordinary *engineer's scale* may prove satisfactory. Engineer's scales, as shown in Fig. 4-2, are available in both metric and English units and have several different graduation intervals so as to accommodate various nominal scales. Precision and accuracy can be enhanced by using a magnifying glass and the finest set of graduations, as long as suitable care is taken in aligning and reading the scale.

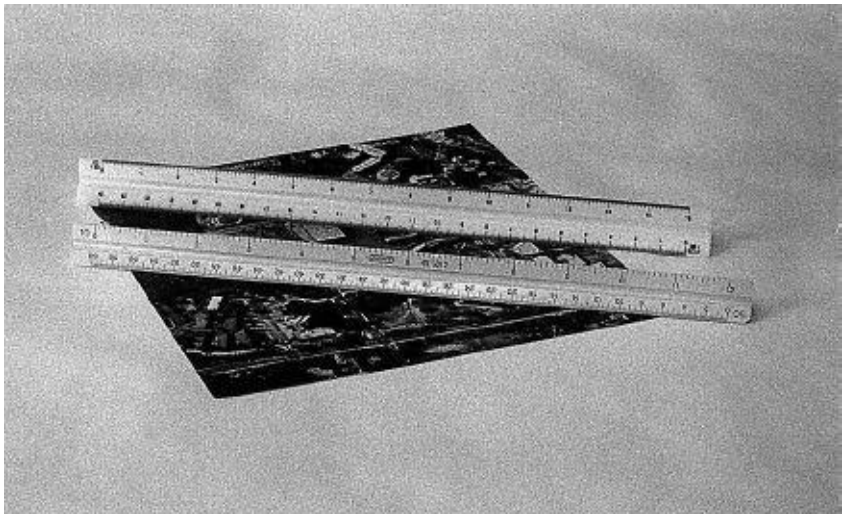


FIGURE 4-2 Metric (top) and English (bottom) engineer's scales.

When greater accuracy is desired, a device such as the glass scale of Fig. 4-3 may be used. Glass scales have fine graduations etched on the bottom surface and are equipped with magnifying eyepieces which can slide along the scale. Glass scales of the type shown in Fig. 4-3 can be obtained in either 6- or 12-in length with either millimeter graduations (least graduations of 0.1 mm) or inch graduations (least graduations of 0.005 in). With a glass scale, readings may be estimated quite readily to one-tenth of the smallest division, but these scales cannot be used to lay off distances.

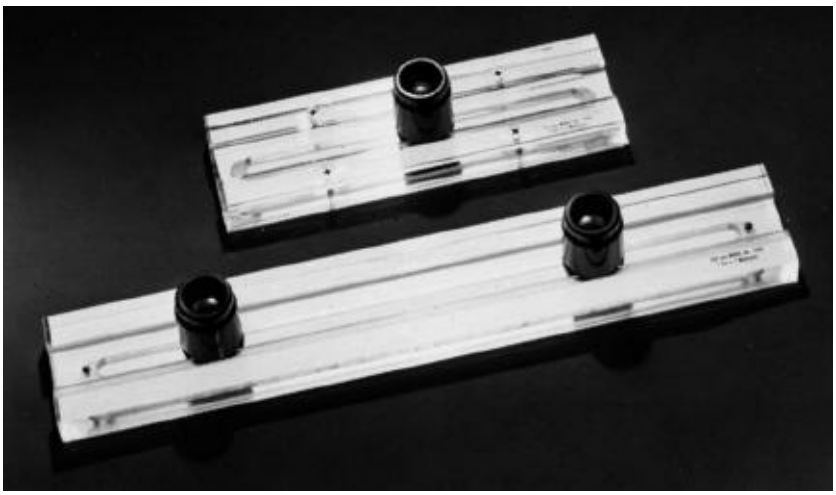


FIGURE 4-3 Glass scales for photographic measurements. (*Courtesy Teledyne-Gurley Co.*)

4-4 Measuring Photo Coordinates with Simple Scales

The conventional procedure for measuring photo coordinates when using an engineer's scale generally consists of first marking the photo coordinate axis system. This may be done by carefully aligning a straightedge across the fiducial marks and lightly making a line with a razor blade, pin, or very sharp 4H or 5H pencil. Rectangular coordinates are then obtained by direct measurement of the perpendicular distances from these axes. Use of glass scales is generally not warranted when this method of marking axes is used, since the enhanced precision and accuracy of the scale are overshadowed by the errors associated with marking the axes.

If the points whose coordinates are to be measured are sharp, distinct points, they may need no further identification. If not, they may be identified with a small pinprick. This should be carefully

done under magnification, however, because systematic error will be introduced into measured photo coordinates if points are erroneously marked.

It is important to affix the proper algebraic sign to measured rectangular coordinates; failure to do so will result in frustrating mistakes in solving photogrammetry problems. Points situated to the right of the y axis have positive x coordinates, and points to the left have negative x coordinates. Points above the x axis have positive y coordinates, and those below the x axis have negative y coordinates.

4-5 Comparator Measurement of Photo Coordinates

For direct measurement of film, the ultimate in photo coordinate measurement accuracy can be achieved with precise instruments called *comparators*. While no longer in common use, these instruments are so named because they compare the photographic positions of imaged points with respect to the measurement scales of the devices. There are two basic types of comparators, *monocomparators* and *stereocomparators*. Monocomparators make measurements on one photograph at a time and with stereocomparators, image positions are measured by simultaneously viewing an overlapping stereo pair of photographs. Comparators are used primarily to obtain precise photo coordinates necessary for camera calibration and for analytical photogrammetry. Their accuracy capability is typically in the range of from 2 to 3 micrometers (μm).

The usual approach for measuring photographic coordinates with a comparator is to measure the coordinates of all image points as well as all fiducial marks. Then an affine or other two-dimensional coordinate transformation is performed in order to relate the arbitrary comparator coordinates to the axis system related to the fiducials. Example C-3. (of App. C) illustrates the procedure using a two-dimensional affine transformation.

Traditional monocomparators are no longer being manufactured today, although specialized devices are still being produced for limited markets. A more versatile device, the *analytical stereoplotter*, is commonly used to perform the function of both monocomparators and stereocomparators when working directly with film. Analytical stereoplotters are described in greater detail in Chap. 12.

Although monocomparators are generally very precise, small systematic errors do occur as a result of imperfections in their measurement systems. The magnitudes of these errors can be determined by measuring coordinates of a precise grid plate and then comparing the results with known coordinate values of the grid plate. The overall pattern of differences (errors) can be modeled with polynomials in a manner similar to that described in Sec. B-11. Measured photo coordinates can then be processed through the polynomial to effectively eliminate the systematic errors of the comparator.

4-6 Photogrammetric Scanners

Photogrammetric scanners are devices used to convert the content of photographs from *analog form* (a continuous-tone image) to *digital form* (an array of pixels with their gray levels quantified by numerical values). These concepts were introduced in Sec. 2-12. Once the image is in digital form, coordinate measurement can take place in a computer environment, either through a manual process involving a human operator who points at features displayed on a computer screen, or through automated image-processing algorithms. A number of photogrammetric quality scanners are available, however their manufacture has greatly declined due to increased use of digital cameras.

They vary in approaches taken in the digital conversion (or *quantization*); however, their fundamental concepts are the same. It is essential that a photogrammetric scanner have sufficient geometric and radiometric resolution as well as high geometric accuracy.

The notions of geometric and radiometric resolution were previously discussed in Sec. 2-12. Geometric or spatial resolution of a scanner is an indication of the pixel size of the resultant image. The smaller the pixel size, the greater the detail that can be detected in the image. High-quality photogrammetric scanners should be capable of producing digital images with minimum pixel sizes on the order of 5 to 15 μm. This roughly corresponds to the resolution threshold of typical aerial photographs under actual flight conditions. Radiometric resolution of a scanner is an indication of the number of quantization levels (corresponding to image density differences) associated with a single pixel. Minimum radiometric resolution should be 256 levels (8-bit) with most scanners being capable of 1024 levels (10-bit) or higher.

The geometric quality of a scanner can be expressed by the positional accuracy of the pixels in the resultant image. If a digital image is to produce the same level of accuracy as is attainable by using film images and a comparator, the positions of the pixels in the digital image need to be at the same spatial accuracy. Hence, the geometric positional accuracy of a high-quality photogrammetric scanner should be at the 2- to 3-μm level. Measurements from scanned photos is similar to those using a comparator. As noted in Sec. 4-5, coordinates of image points and fiducial marks are measured, but in this case as rows and columns in the scanner coordinate system. Then a two-dimensional coordinate transformation is performed as in Example C-3, to produce photo coordinates related to the fiducials.

4-7 Refinement of Measured Image Coordinates

The preceding sections of this chapter have discussed instruments and techniques for measuring photo coordinates. Procedures have also been described for eliminating systematic errors in the measurements and for reducing the coordinates to the photo coordinate axis system. These photo coordinates will still contain systematic errors from various other sources, however. The major sources of these errors are

1. (a) Film distortions due to shrinkage, expansion, and lack of flatness; (b) CCD array distortions due to electrical signal timing issues or lack of flatness of the chip surface
2. (a) Failure of photo coordinate axes to intersect at the principal point; (b) Failure of principal point to be aligned with center of CCD array
3. Lens distortions
4. Atmospheric refraction distortions
5. Earth curvature distortion

Corrections may be applied to eliminate the effects of these systematic errors. However, not all corrections need to be made for all photogrammetric problems; in fact, for work of coarse accuracy they may all be ignored. If, for example, an engineer's scale has been used to make the measurements, uncertainty in the photo coordinates may be so great that the small magnitudes of these systematic errors become insignificant. On the other hand, if precise measurements for an analytical photogrammetry problem have been made with a comparator, all the corrections may be significant. The decision as to which corrections are necessary for a particular photogrammetric problem can be made after considering required accuracy versus magnitude of error caused by neglecting the correction.

4-8 Distortions of Photographic Films and Papers

In photogrammetric work, true positions of images in the picture are required. Photo coordinates measured by any of the previously discussed methods will unavoidably contain small errors due to shrinkage or expansion of the photographic materials that support the emulsion of the negative and positive. In addition, since photogrammetric equations derived for applications involving frame cameras assume a flat image plane, any lack of flatness will likewise cause errors. These errors can also be categorized as film distortions, and they are generally the most difficult to compensate for, due to their nonhomogeneous nature. Photo coordinates must be corrected for these errors before they are used in photogrammetric calculations; otherwise, errors will be present in the computed results. The magnitude of error in computed values will depend upon the severity of the film distortions, which depends upon the type of emulsion support materials used and the flatness of the camera platen.

Most photographic films used to produce negatives for photogrammetric work have excellent dimensional stability, but some small changes in size do occur during processing and storage. Dimensional change during storage may be held to a minimum by maintaining constant temperature and humidity in the storage room. The actual amount of distortion present in a film is a function of several variables, including the type of film and its thickness. Typical values may vary from almost negligible amounts up to approximately 0.2 percent.

Paper media are generally much less stable than film. Whether the images are photographically reproduced or are printed from a scanned image, paper copies often require corrections to measured distances in order to produce satisfactory results.

4-9 Image Plane Distortion

The nominal amount of shrinkage or expansion present in a photograph can be determined by comparing measured photographic distances between opposite fiducial marks with their corresponding values determined in camera calibration. Photo coordinates can be corrected if discrepancies exist, and the approach differs depending on the necessary level of accuracy. For lower levels of accuracy (corresponding to measurements with an engineer's scale on paper prints) the following approach may be used. If x_m and y_m are measured fiducial distances on the positive, and x_c and y_c are corresponding calibrated fiducial distances, then the corrected photo coordinates of any point a may be calculated as

$$x'_a = \left(\frac{x_c}{x_m}\right) x_a \tag{4-2}$$

$$y'_a = \left(\frac{y_c}{y_m}\right) y_a \tag{4-3}$$

In Eqs. (4-2) and (4-3), x'_a and y'_a are corrected photo coordinates and x_a and y_a are measured coordinates. The ratios x_c/x_m and y_c/y_m are simply scale factors in the x and y directions, respectively. This method is also appropriate for images that have been subjected to substantial enlargement or are being measured with arbitrary coordinates, for example, an image from a digital camera having a CCD frame width of 20 mm printed at a width of 150 mm on a sheet of paper. If the precise dimensions of the CCD frame have been determined through calibration, a correction factor based on Eq. (4-2) can

be computed. In another case, a digital image may be imported into a Computer Aided Drafting (CAD) drawing. In this case the coordinates of the CAD drawing may be completely arbitrary and need to be related to a calibrated reference in order to obtain correct measurements. Equations (4-2) and (4-3) can be used by making CAD-unit measurements of the calibrated distances and applying the correction factors to subsequent measurements.

Example 4-1
For a particular photograph, the measured x and y fiducial distances were 233.8 and 233.5 mm, respectively. The corresponding x and y calibrated fiducial distances were 232.604 and 232.621 mm, respectively. Compute the corrected values for the measured photo coordinates which are listed in columns (*b*) and (*c*) in the table below.

	Measured Coordinates		Corrected Coordinates	
(*a*) Point no.	(*b*) x, mm	(*c*) y, mm	(*d*) x', mm	(*e*) y', mm
1	−102.6	95.2	−102.1	94.8
2	−98.4	−87.8	−97.9	−87.5
3	16.3	−36.1	16.2	−36.0
4	65.7	61.8	65.4	61.6
5	104.9	−73.5	104.4	−73.2

Solution From Eq. (4-2),

$$x' = \frac{232.604}{233.8} x = 0.9949x$$

$$y' = \frac{232.621}{233.5} y = 0.9962y$$

Each of the measured values is multiplied by the appropriate constant above, and the corrected coordinates are entered in columns (*d*) and (*e*) of the table. ▲

For high-accuracy applications, shrinkage or expansion corrections may be applied through the x and y scale factors of a two-dimensional affine coordinate transformation. This method is particularly well-suited for analytical photogrammetric calculations. This procedure is described in Sec. C-6, and a numerical example is presented.

4-10 Reduction of Coordinates to an Origin at the Principal Point

It has been previously stated that the principal point of a photograph rarely occurs precisely at the intersection of fiducial lines or at the center of a CCD array. The actual coordinates of the principal point with respect to the calibrated xy photo coordinate system of the camera are x_p and y_p, shown in Fig. 3-18. These coordinates specify the principal point location about which the lens distortions are most symmetric.

Photogrammetric equations that utilize photo coordinates are based on projective geometry and

assume an origin of photo coordinates at the principal point. Therefore it is theoretically correct to reduce photo coordinates from the measurement or photo coordinate axis system to the axis system whose origin is at the principal point. Manufacturers of precision mapping cameras attempt to mount the fiducial marks and camera lens so that the principal point and intersection of fiducial lines coincide. Normally they accomplish this to within a few micrometers, and therefore in work of coarse accuracy using engineering scales and paper prints, this correction can usually be ignored. For precise analytical photogrammetric work, it is necessary to make the correction for the coordinates of the principal point. The correction is applied after a two-dimensional coordinate transformation (e.g., affine) is made to the coordinates measured by comparator or from a scanned image. The principal point coordinates x_p and y_p from the camera calibration report are subtracted from the transformed x and y coordinates, respectively. Most appropriately, the correction for the principal point offset is applied in conjunction with lens distortion corrections.

4-11 Correction for Lens Distortions

As described in Chap. 2, lens distortion causes imaged positions to be displaced from their ideal locations. The mathematical equations that are used to model lens distortions are typically comprised of two components: symmetric radial distortion and decentering distortion. In modern precision aerial mapping cameras, lens distortions are typically less than 5 μm and are only applied when precise analytical photogrammetry is being performed.

Symmetric radial lens distortion is an unavoidable product of lens manufacture, although with careful design its effects can be reduced to a very small amount. Decentering distortion, on the other hand, is primarily a function of the imperfect assembly of lens elements, not the actual design. Historically, metric aerial mapping cameras had significantly larger amounts of symmetric radial lens distortion than decentering distortion. Traditional camera calibration procedures provided information regarding only the symmetric radial component. For instance, an early approach used by the U.S. Geological Survey was to indicate radial distortion values for each of the angles of the multicollimator (see Sec. 3-11). The radial distortion value was the radial displacement from the ideal location to the actual image of the collimator cross, with positive values indicating outward displacements. The approach used for determining radial lens distortion values for these older calibration reports was to fit a polynomial curve to a plot of the displacements (on the ordinate) versus radial distances (on the abscissa). The form of the polynomial, based on lens design theory, is

$$\Delta r = k_1 r^1 + k_2 r^3 + k_3 r^5 + k_4 r^7 \tag{4-4}$$

In Eq. (4-4), Δr is the amount of radial lens distortion, r is the radial distance from the principal point, and k_1, k_2, k_3, and k_4 are coefficients of the polynomial. The coefficients of the polynomial are solved by least squares using the distortion values from the calibration report. To correct the x, y position of an image point, the distance r from the image point to the principal point is computed and used to compute the value of Δr from Eq. (4-4). This is done by first converting the fiducial coordinates x and y, to coordinates $\bar{x}$ and $\bar{y}$, relative to the principal point, by Eqs. (4-5) and (4-6). Then Eq. (4-7) is used to compute the value of r to use in Eq. (4-4).

$$\bar{x} = x - x_p \tag{4-5}$$

$$\bar{y} = y - y_p \tag{4-6}$$

$$r = \sqrt{\bar{x}^2 + \bar{y}^2} \tag{4-7}$$

After the radial lens distortion value of Δr is computed, its x and y components (corrections δx and δy) are computed and subtracted from $\bar{x}$ and $\bar{y}$, respectively. The δx and δy corrections are based on a similar-triangle relationship, as shown in Fig. 4-4. By similar triangles of that figure

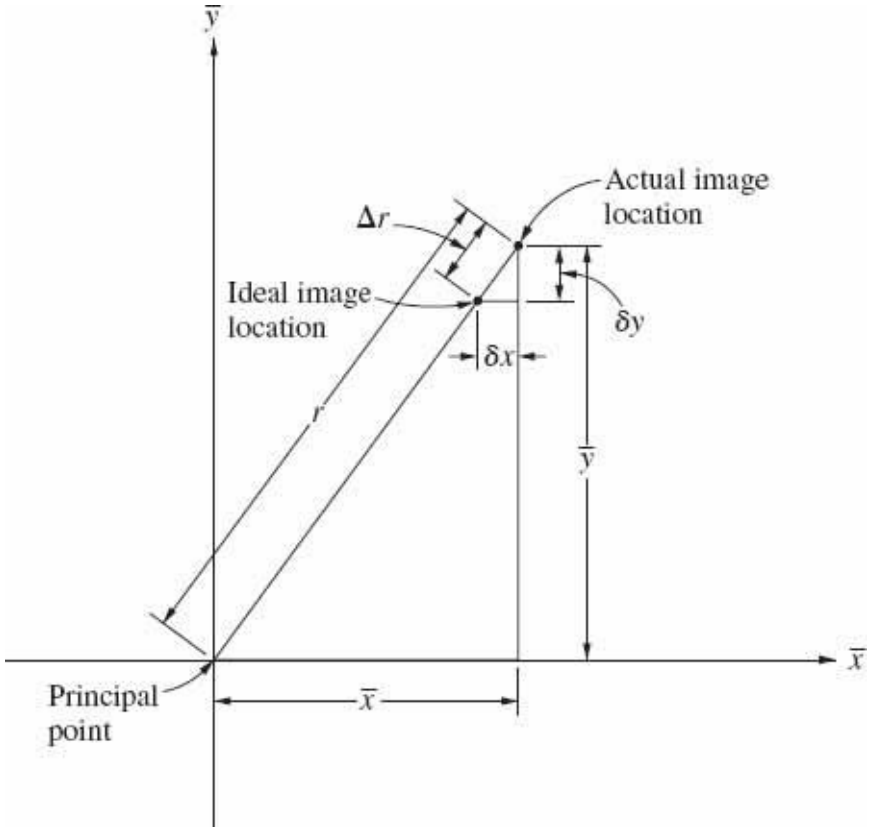


FIGURE 4-4 Relationship between radial lens distortion and corrections to x and y coordinates.

$$\frac{\Delta r}{r} = \frac{\delta x}{\bar{x}} = \frac{\delta y}{\bar{y}}$$

from which

$$\delta x = \bar{x}\frac{\Delta r}{r} \tag{4-8}$$

$$\delta y = \bar{y}\frac{\Delta r}{r} \tag{4-9}$$

The corrected coordinates x_c and y_c are then computed by

$$x_c = \bar{x} - \delta x \tag{4-10}$$

$$y_c = \bar{y} - \delta y \tag{4-11}$$

Example 4-2

An older USGS camera calibration report specifies the calibrated focal length f = 153.206 mm and coordinates of the calibrated principal point as x_p = 0.008 mm and y_p = –0.001 mm. The report also lists mean radial lens distortion values given in columns (*a*) and (*b*) of the table below. Using these calibration values, compute the corrected coordinates for an image point having coordinates x = 62.579 mm, y = –80.916 mm relative to the fiducial axes.

Mean Radial Distortion		
(a) Field angle θ	(b) Δr, mm	(c) r, m
7.5°	0.004	0.0202
15°	0.007	0.0411
22.5°	0.007	0.0635
30°	0.001	0.0885
35°	–0.003	0.1073
40°	–0.004	0.1286

Solution Compute r values (in meters) in column (*c*) by the following equation (see Fig. 4-5):

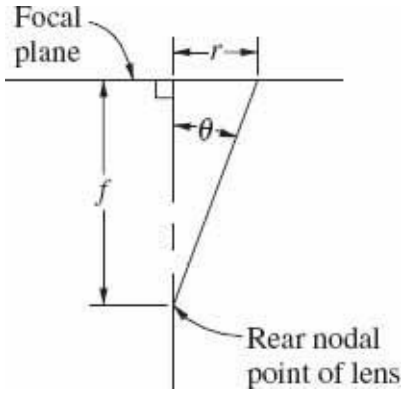


FIGURE 4-5 Illustration of radial distance r as it relates to focal length f and field angle θ.

$$r = f \tan \theta$$

For example, for the field angle, θ = 7.5°,

$$r = 0.153206 \tan 7.5° = 0.0202 \text{ m}$$

Using the least squares method presented in Sec. B-11 (see Example B-6), the following k values were computed.

$$k_1 = 0.2296 \qquad k_3 = 1018$$

$$k_2 = -35.89 \qquad k_4 = 12{,}100$$

Compute the distance from the principal point to the image point, using Eqs. (4-5), (4-6), and (4-7).

$$\bar{x} = x - x_p = 62.579 - 0.008 = 62.571 \text{ mm} = 0.062571 \text{ m}$$

$$\bar{y} = y - y_p = -80.916 - (-0.001) = -80.915 \text{ mm} = -0.080915 \text{ m}$$

$$r = \sqrt{0.062571^2 + (-0.080915)^2} = 0.1023 \text{ m}$$

Given this value for r and the k coefficients, compute Δr by Eq. (4-4):

$$\Delta r = (0.2296)(0.1023) + (-35.89)(0.1023)^3 + (1018)(0.1023)^5 + (12{,}100)(0.1023)^7$$

$$= -0.0021 \text{ mm}$$

Compute δx and δy by Eqs. (4-8) and (4-9), respectively.

$$\delta x = 0.062571\left(\frac{-0.0021}{0.1023}\right) = -0.0013 \text{ mm}$$

$$\delta y = -0.080915\left(\frac{-0.0021}{0.1023}\right) = 0.0017 \text{ mm}$$

Compute corrected coordinates x_c and y_c by Eqs. (4-10) and (4-11), respectively.

$$x_c = 62.571 - (-0.0013) = 62.572 \text{ mm}$$

$$y_c = -80.915 - 0.0017 = -80.917 \text{ mm} \quad \blacktriangle$$

Lens design in modern aerial mapping cameras has evolved to such a level that symmetric radial lens distortion is of the same order of magnitude as decentering distortion, and camera calibration reports have been adapted to accommodate this change. For example, the mathematical model used in the current USGS calibration procedure, known as the Simultaneous Multi-camera Analytical Calibration (SMAC), computes both symmetric radial and decentering distortion parameters directly by least squares. Principal point coordinates and focal length are also determined in the solution. The USGS camera calibration report lists polynomial coefficients for symmetric radial lens distortion (k_0, k_1, k_2, k_3, k_4), and decentering distortion (p_1, p_2, p_3, p_4). It also gives calibrated principal point coordinates (x_p, y_p). To compute coordinates (x_c, y_c) corrected for these systematic errors, the following equations are used:

$$\delta x = \bar{x}(k_0 + k_1 r^2 + k_2 r^4 + k_3 r^6 + k_4 r^8) \tag{4-12}$$

$$\delta y = \bar{y}(k_0 + k_1 r^2 + k_2 r^4 + k_3 r^6 + k_4 r^8) \tag{4-13}$$

$$\Delta x = (1 + p_3 r^2 + p_4 r^4)[p_1(r^2 + 2\bar{x}^6) + 2p_2 \bar{x}\,\bar{y}] \tag{4-14}$$

$$\Delta y = (1 + p_3 r^2 + p_4 r^4)[2p_1 \bar{x}\,\bar{y} + p_2(r^2 + 2\bar{y}^2)] \tag{4-15}$$

$$x_c = \bar{x} + \delta x + \Delta x \tag{4-16}$$

$$y_c = \bar{y} + \delta y + \Delta y \tag{4-17}$$

In Eqs. (4-12) through (4-17), $\bar{x}$ and $\bar{y}$ are coordinates of the image relative to the principal point as computed by Eqs. (4-5) and (4-6), respectively; r is the radial distance from the image to the principal point as computed by Eq. (4-7); k_0, k_1, k_2, k_3, and k_4 are coefficients of symmetric radial lens distortion from the calibration report; p_1, p_2, p_3, and p_4 are coefficients of decentering distortion from the calibration report; δx and δy are the symmetric radial lens distortion corrections to $\bar{x}$ and $\bar{y}$, respectively; and Δx and Δy are the decentering distortion corrections to $\bar{x}$ and $\bar{y}$, respectively.

Example 4-3

The parameters of a current USGS camera calibration report are given in the following table. Using these calibration values, compute the corrected coordinates for an image point having coordinates x = –47.018 mm, y = 43.430 mm relative to the fiducial axes.

Symmetric Radial Distortion Parameters		Decentering Distortion Parameters		Calibrated Principal Point	
k_0	0.5493×10^{-4}	p_1	-0.7953×10^{-7} mm^{-1}	x_p	0.010 mm
k_1	-0.5984×10^{-8} mm^{-2}	p_2	0.1018×10^{-6} mm^{-1}	y_p	–0.001 mm
k_2	0.1053×10^{-12} mm^{-4}	p_3	0 mm^{-2}		
k_3	0 mm^{-6}	p_4	0 mm^{-4}		
k_4	0 mm^{-8}				

Solution Compute $\bar{x}$, $\bar{y}$, and r by Eqs. (4-5), (4-6), and (4-7), respectively.

$$\bar{x} = -47.018 - 0.010 = -47.028 \text{ mm}$$

$$\bar{y} = 43.430 - (-0.001) = 43.431 \text{ mm}$$

$$r = \sqrt{(-47.028)^2 + 43.431^2} = 64.015 \text{ mm}$$

Compute symmetric radial lens distortion corrections δx and δy, using Eqs. (4-12) and (4-13),

respectively.

$$\delta x = -47.028[0.5493 \times 10^{-4} - 0.5984 \times 10^{-8}(64.015)^2 + 0.1053 \times 10^{-12}(64.015)^4] = -0.0015 \text{ mm}$$

$$\delta y = 43.431[0.5493 \times 10^{-4} - 0.5984 \times 10^{-8}(64.015)^2 + 0.1053 \times 10^{-12}(64.015)^4] = 0.0014 \text{ mm}$$

Compute decentering distortion corrections Δx and Δy, using Eqs. (4-14) and (4-15), respectively.

$$\Delta x = 1\{(-0.7953 \times 10^{-7})[64.015^2 + 2(-47.028)^2] + 2(0.1018 \times 10^{-6})(-47.028)(43.431)\} = -0.0011 \text{ mm}$$

$$\Delta y = 1\{2(-0.7953 \times 10^{-7})(-47.028)(43.431) + (0.1018 \times 10^{-6})[64.015^2 + 2(43.431)^2]\} = 0.0011 \text{ mm}$$

Compute the corrected coordinates x_c and y_c, using Eqs. (4-16) and (4-17), respectively.

$$x_c = -47.028 - 0.0015 - 0.0011 = -47.031 \text{ mm}$$

$$y_c = 43.431 + 0.0014 + 0.0011 = 43.434 \text{ mm} \quad \blacktriangle$$

4-12 Correction for Atmospheric Refraction

It is well known that density (and hence the index of refraction) of the atmosphere decreases with increasing altitude. Because of this condition, light rays do not travel in straight lines through the atmosphere, but rather they are bent according to Snell's law (see Sec. 2-2), as shown in Fig. 4-6. The incoming light ray from point *A* of the figure makes an angle α with the vertical. If refraction were ignored, the light ray would appear to be coming from point *B* rather than from point *A*. Photogrammetric equations assume that light rays travel in straight paths, and to compensate for the known refracted paths, corrections are applied to the image coordinates.

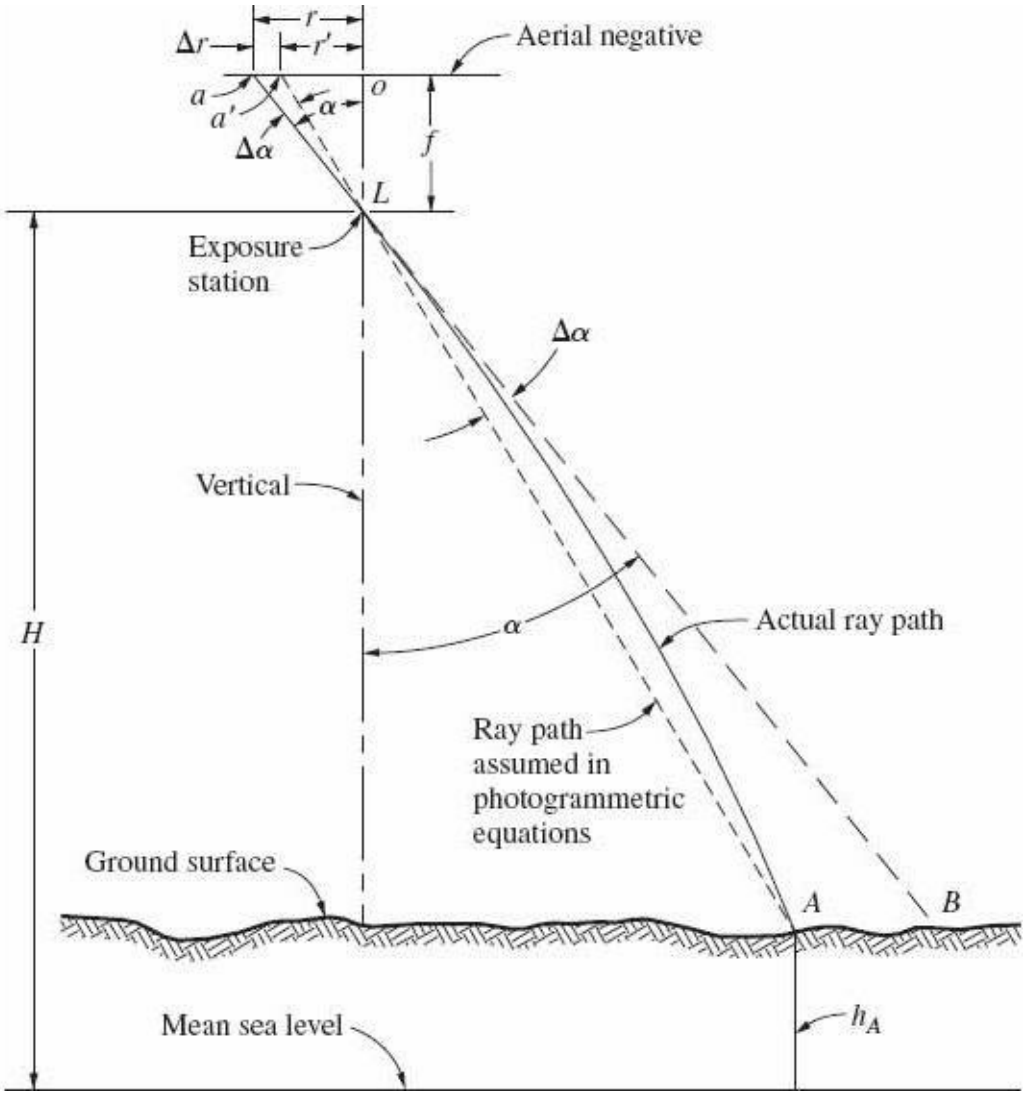


FIGURE 4-6 Atmospheric refraction in aerial photography.

In Fig. 4-6, if a straight path had been followed by the light ray from object point *A*, then its image would have been at *a'*. The angular distortion due to refraction is $\Delta\alpha$, and the linear distortion on the photograph is Δr. Refraction causes all imaged points to be displaced outward from their correct positions. The magnitude of refraction distortion increases with increasing flying height and with increasing α angle. Refraction distortion occurs radially from the photographic nadir point[1] (principal point of a vertical photo) and is zero at the nadir point. Atmospheric refraction in tilted photographs is treated in Sec. 10-16.

The usual approach to the atmospheric refraction correction is based on the assumption that change in the refractive index of air is directly proportional to change in height. Starting with the incident angle of the light ray at ground level, Snell's law can be solved continuously along the ray path for each infinitesimal change in angle due to refraction. When all the infinitesimal changes are summed, the total is proportional to the tangent of the incident angle. (Derivation of the relationship involves the solution of a differential equation, which is beyond the scope of this text.) The proportionality constant is based on the values of the refractive indices at ground level and at the camera position, which are related to elevation. The relationship that expresses the angular distortion $\Delta\alpha$ as a function of α is

$$\Delta\alpha = K \tan \alpha \tag{4-18}$$

In this equation, α is the angle between the vertical and the ray of light, as shown in Fig. 4-6, and K is a value which depends upon the flying height above mean sea level and the elevation of the object point. There are several different approaches to calculating a value for K, with most assuming a standard atmosphere. A convenient method, adapted from the *Manual of Photogrammetry*, is to compute K by

$$K = (7.4 \times 10^{-4})\,(H - h)\,[1 - 0.02(2H - h)] \tag{4-19}$$

In Eq. (4-19), H is the flying height of the camera above mean sea level in kilometers, and h is the elevation of the object point above mean sea level in kilometers. The units of K are degrees.

The procedure for computing atmospheric refraction corrections to image coordinates on a vertical photo begins by computing radial distance r from the principal point to the image, using Eq. (4-20). In this equation, the x and y image coordinates do not necessarily need to be related to the principal point since the error due to the assumption of vertical photography far overshadows any error which would be introduced.

$$r = \sqrt{r^2 + y^2} \tag{4-20}$$

Also from Fig. 4-6,

$$\tan\alpha = \frac{r}{f} \tag{4-21}$$

The values of K and tan α from Eqs. (4-19) and (4-21), respectively, are then substituted into Eq. (4-18) to compute refraction angle $\Delta\alpha$.

$$\Delta\alpha = K\frac{r}{f} \tag{4-22}$$

The radial distance r' from the principal point to the corrected image location can then be computed by

$$r' = f\tan(\alpha - \Delta\alpha) \tag{4-23}$$

The change in radial distance Δr is then computed by

$$\Delta r = r - r' \tag{4-24}$$

The x and y components of atmospheric refraction distortion corrections (δx and δy) can then be computed by Eqs. (4-8) and (4-9), using the values of x and y in place of $\bar{x}$ and $\bar{y}$, respectively. To compute corrected coordinates x' and y', the corrections δx and δy are subtracted from x and y, respectively.

Example 4-4

A vertical photograph taken from a flying height of 3500 m above mean sea level contains the image a of object point A at coordinates (with respect to the fiducial system) x_a = 73.287 mm and y_a = –101.307 mm. If the elevation of object point A is 120 m above mean sea level and the camera had a focal length of 153.099 mm, compute the x' and y' coordinates of the point, corrected for atmospheric refraction.

Solution Compute r by Eq. (4-20).

$$r = \sqrt{73.287^2 + (-101.307)^2} = 125.036 \text{ mm}$$

Express tan α by Eq. (4-21) and solve for α.

$$\tan\alpha = \frac{125.036}{153.099}$$

$$\alpha = 39.2386°$$

Compute K by Eq. (4-19).

$$K = (7.4 \times 10^{-4})\,(3.5 - 0.12)\,\{1 - 0.02[2(3.5) - 0.12]\} = 0.0022°$$

Compute $\Delta\alpha$ by Eq. (4-22).

$$\Delta\alpha = 0.0022° \left(\frac{125.036}{153.099}\right) = 0.0018°$$

Compute r' by Eq. (4-23).

$$r' = 153.099 \tan (39.2386° - 0.0018°) = 125.029 \text{ mm}$$

Compute Δr by Eq. (4-24).

$$\Delta r = 125.036 - 125.029 = 0.008 \text{ mm}$$

Compute δx and δy by Eqs. (4-8) and (4-9), respectively.

$$\delta x = 73.287 \left(\frac{0.008}{125.036}\right) = 0.005 \text{ mm}$$

$$\delta y = -101.307 \left(\frac{0.008}{125.036}\right) = -0.006 \text{ mm}$$

Subtract the corrections δx and δy from x and y, respectively to obtain corrected coordinates x' and y'.

$$x' = x - \delta x = 73.287 - 0.005 = 73.282 \text{ mm}$$

$$y' = y - \delta y = -101.307 - (-0.006) = -101.301 \text{ mm} \quad ▲$$

4-13 Correction for Earth Curvature

Traditionally, in analytical photogrammetry, corrections were commonly applied to measured photo coordinates to compensate for the effects of earth curvature. The rationale for this notion is that elevations of points are referenced to an approximately spherical datum (i.e., mean sea level) whereas photogrammetric equations assume that the zero-elevation surface is a plane. In addition, if horizontal coordinate systems such as state plane coordinates (see Sec. 5-6) are used in object space, the axis system is also curved, although it is curved primarily in only one direction (either east-west or north-south). It has long been recognized that the practice of making earth curvature corrections to measured photo coordinates is not theoretically correct. However, it has also long been accepted that using the correction generally leads to more accurate results than ignoring it, particularly in determining elevations.

The primary problem with the earth curvature correction is that because of the nature of map projection coordinates, correcting photo coordinates for earth curvature will degrade the accuracy of either *X* or *Y* object space coordinates, depending upon the map projection used. For example, when a UTM projection (see Sec. 5-6) is used, application of earth curvature corrections will yield more accurate elevations and *Y* values, but will degrade *X* values because the UTM projection does not curve in the *X* direction.

The proper approach, which avoids the need for any sort of earth curvature correction, employs a three-dimensional orthogonal object space coordinate system. One such coordinate system is the *local vertical coordinate system* described in Sec. 5-5. By using a local vertical coordinate system for a project, earth curvature ceases to be a distorting effect. Instead, the curvature of the earth will simply be a natural characteristic of the terrain, and it will be resolved in the same manner as any other topographic feature.

4-14 Measurement of Feature Positions and Edges

Up to this point in the chapter, measurement of image distances and coordinates of points have been discussed in a general context. Now we will consider some of the characteristics of features to be measured, namely points and edges. Since a “point” must be visible on an image in order to be measured, the feature must have finite size. Small round objects such as manhole covers will appear as nearly ideal point features as long as the scale of the image is not too large. Under magnification or image zoom, the edges of the object will appear to be somewhat indistinct although the center can still be identified to a high precision as shown in Fig. 4-7. If this manhole appears in another image, its center will be a repeatable measurement point whereas its edges may not.

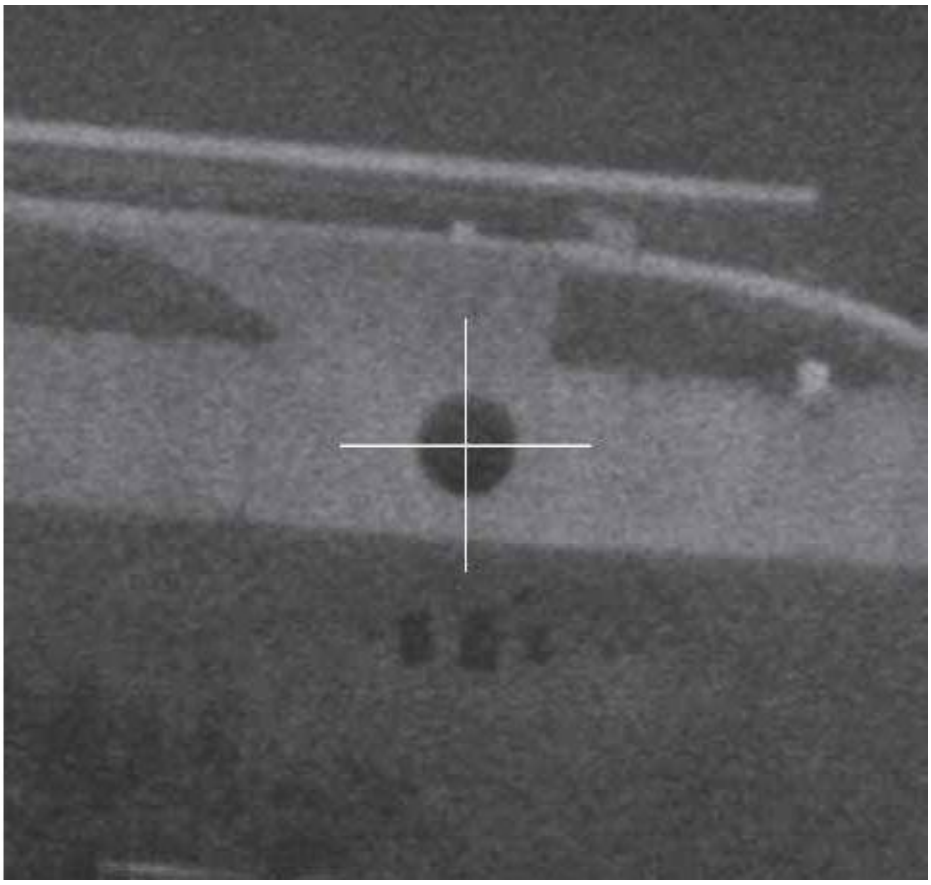


FIGURE 4-7 Zoomed-in view of a manhole with cross-hair at center.

In a situation where a point feature is designated to be at the angle point of a painted chevron target, it is important to use the intersections of the centerlines of the stripes rather than the outside or inside corner as the point. Edges that may appear clear and sharp at a glance will become blurred and indistinct when viewed under magnification as shown in Fig. 4-8. In this figure, it would be difficult to precisely place cross-hairs on the corner of the chevron, but the centerlines of the stripes can be identified precisely.

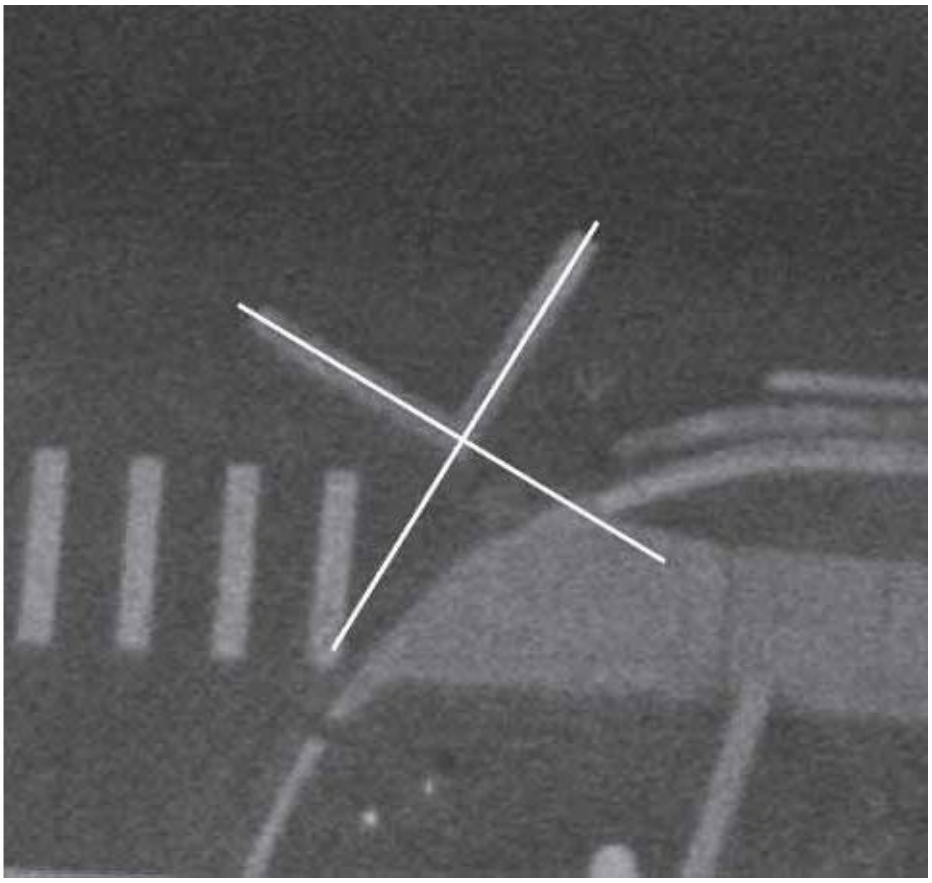


FIGURE 4-8 Zoomed-in view of a chevron target with cross-hair at the intersection of centerlines of stripes.

In many cases it is necessary to measure edges, particularly when mapping features in an area. It is then important to make a best estimate of where the true edge of the feature is located. A certain amount of error can be expected in the measurement due to blurred edges. For example, consider measuring the lengths of the stripes of the chevron in Fig. 4-8. It is very difficult to identify the precise location of where the paint meets the underlying pavement.

References

Abdel-Aziz, Y. L.: "Asymmetrical Lens Distortion," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 3, 1975, p. 337.

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004, chaps. 3 and 10.

Brown, D. C.: "Close-Range Camera Calibration," *Photogrammetric Engineering*, vol. 37, no. 8, 1971, p. 855.

———: "Unflatness of Plates as a Source of Systematic Error in Close-Range Photogrammetry," *Photogrammetria*, vol. 40, no. 9, 1986, p. 343.

Forrest, R. B.: "Refraction Compensation," *Photogrammetric Engineering*, vol. 40, no. 5, 1974, p. 577.

Fraser, C. S.: "Photogrammetric Measurement to One Part in a Million," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 3, 1992, p. 305.

Fritz, L. W.: "A Complete Comparator Calibration Program," *Photogrammetria*, vol. 29, no. 4, 1973, p. 133.

Frost, R. M.: "Improved Well Positions for Geoscientific Applications: Exploiting NAPP Photographs with Digitizer and PC-Based Bundle Adjustment Program," *Photogrammetric Engineering and*

Remote Sensing, vol. 61, no. 7, 1995, p. 927.
Gyer, M. S.: "Methods for Computing Photogrammetric Refraction Corrections for Vertical and Oblique Photographs," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 3, 1996, p. 301.
Jaksic, Z.: "Deformations of Estar-Base Aerial Films," *Photogrammetric Engineering*, vol. 38, no. 3, 1972, p. 285.
Light, D. L.: "The New Camera Calibration System at the U.S. Geological Survey," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 2, 1992, p. 185.
Murase, T., M. Tanaka, T. Tani, Y. Miyashita, N. Ohkawa, S. Ishiguro, Y. Suzuki, H. Kayanne, and H. Yamano: "A Photogrammetric Correction Procedure for Light Refraction Effects at a Two-Medium Boundary," *Photogrammetric Engineering and Remote Sensing*, vol. 74, no. 9, 2008, p. 1129.
Scarpace, F. L., and P. R. Wolf: "Atmospheric Refraction," *Photogrammetric Engineering*, vol. 39, no. 5, 1973, p. 521.
Scholer, H.: "On Photogrammetric Distortion," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 6, 1975, p. 761.
Tayman, W. P.: "User Guide for the USGS Aerial Camera Report of Calibration," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 5, 1984, p. 577.
Wiley, A. G., and K. W. Wong: "Geometric Calibration of Zoom Lenses for Computer Vision Metrology," *Photogrammetric Engineering and Remote Sensing*, vol. 61, no. 1, 1995, p. 69.

Problems

4-1. Assume that photo coordinates of points *a* and *b* of Fig. 4-1 are $x_a = 57.62$ mm, $y_a = 67.15$ mm, $x_b = 22.07$ mm, and $y_b = -49.50$ mm. Calculate photo distance *ab* and radial distances *oa* and *ob*.

4-2. Repeat Prob. 4-1 except that the photo coordinates are $x_a = -71.66$ mm, $y_a = 47.82$ mm, $x_b = 81.43$ mm, and $y_b = 6.05$ mm.

4-3. Name and briefly describe the various systematic errors that may exist in photographic coordinates.

4-4. Calculate the acute angle (to the nearest 0.001°) of intersection of fiducial lines for a camera of the type shown in Fig. 4-1 if comparator measurements of the fiducial marks on a calibration "flash plate" were as follows:

Mark	*X*, mm	*Y*, mm
1	13.729	128.568
2	126.653	15.652
3	239.549	128.577
4	126.662	241.440

4-5. Repeat Prob. 4-4, except that the following flash plate measurements were taken:

Mark	X, mm	Y, mm
1	8.635	122.458
2	123.007	9.574
3	236.396	122.476
4	123.022	235.347

4-6. If the intersection of fiducial lines of the camera of Prob. 4-4 defines the principal point exactly, what are the x and y photo coordinates of the four fiducial marks in the photo system? Assume that the x and y photo coordinate axes are parallel to the comparator axes.

4-7. Repeat Prob. 4-6, except that it applies to the data of Prob. 4-5.

4-8. On a paper-print positive, the measured x distance between fiducials (1 and 3) was 224.9 mm and y between fiducials (2 and 4) was 223.5 mm. These x and y distances determined in camera calibration were 225.433 and 225.693 mm, respectively. Using the method of Example 4-1, calculate shrinkage-corrected coordinates of points a, b, and c whose coordinates were measured on the paper print as follows:

Point	x, mm	y, mm
a	20.3	–92.1
b	48.6	85.8
c	–111.1	–102.5

4-9. Repeat Prob. 4-8, except that the measured x distance on a paper-print positive between fiducials 1 and 3 was 225.64 mm and y between fiducials 2 and 4 was 225.83 mm; the calibrated distances between these same fiducials were 224.282 and 224.155 mm, respectively; and measured photo coordinates of points a, b, and c were as follows:

Point	x, mm	y, mm
a	14.63	–75.69
b	–78.08	41.81
c	–104.78	–58.42

4-10. A recent USGS camera calibration report yielded the parameters given in the following table. Using these calibration values, compute the corrected coordinates (to the nearest micrometer) for an image point having coordinates x = 73.952 mm and y = 98.155 mm relative to the fiducial axes.

Symmetric Radial Distortion Parameters		Decentering Distortion Parameters		Calibrated Principal Point	
k_0	-0.1220×10^{-3}	p_1	0.1147×10^{-6} mm^{-1}	x_p	0.007 mm
k_1	0.8255×10^{-8} mm^{-2}	p_2	0.5982×10^{-8} mm^{-1}	y_p	–0.006 mm
k_2	-0.1219×10^{-12} mm^{-4}	p_3	0 mm^{-2}		
k_3	0 mm^{-6}	p_4	0 mm^{-4}		
k_4	0 mm^{-8}				

4-11. Repeat Prob. 4-10, except that the coordinates, with respect to the fiducial system, of the image point are $x = -78.090$ mm and $y = -57.688$ mm.

4-12. The photo coordinates listed below have been corrected for film and lens distortions. The camera that took the photography had a focal length of 151.942 mm, the flying height above mean sea level (MSL) was 8530 m, and the average elevation of the terrain was 310 m above MSL. Calculate the photo coordinates (to the nearest micrometer) corrected for atmospheric refraction using the method of Example 4-4.

Point	*x*, mm	*y*, mm
a	28.738	49.211
b	57.820	–93.705
c	–117.232	–102.794

4-13. Repeat Prob. 4-12, except that the camera lens had a focal length of 88.916 mm, flying height above MSL was 4250 m, the average terrain elevation was 170 m above MSL, and the photo coordinates were as follows:

Point	*x*, mm	*y*, mm
a	59.238	74.281
b	–63.970	–113.444
c	103.296	–98.730

[1] The photographic nadir point is defined in Sec. 10-3.

CHAPTER 5

Object Space Coordinate Systems

5-1 Introduction

Coordinate systems are a fundamental concept associated with spatial data. In the previous chapter, two-dimensional *image space* coordinates were discussed. In this chapter, three-dimensional *object space* coordinate systems are presented and described. Object space coordinate systems have always been important for specifying the relative positions of points in surveying, photogrammetry, and mapping. However, they have recently taken on added significance with the emergence and increasing importance of geographic information systems (GISs), which are heavily dependent upon coordinated data for their function.

Object space in photogrammetry refers to the three-dimensional region that encompasses the physical features imaged in the photographs. Most often object space relates to a portion of the earth's terrain and the natural and cultural features thereon; but it also can relate to other items such as celestial bodies, medical subjects, industrial machines, archeological objects, and many others. When mapping the earth's terrain and natural and cultural features, it is important that all mapped objects be accurately located with respect to an accepted geographic frame of reference. This is particularly important when spatial data from multiple sources are being integrated. If any of the spatial data sets are not accurately defined in an accepted frame of reference, then gaps, overlaps, and mismatches will occur. Several accepted reference systems exist: *geodetic*, *geocentric*, *local vertical*, and *map projection*. General concepts of these reference systems are discussed in this chapter, while mathematical equations which express the relationships between them are given in App. F.

5-2 Concepts of Geodesy

To a person standing in the middle of a prairie, the earth appears to be flat, and within a limited area, it is approximately so. Within this limited range, positions of points can be determined relative to a local plane surface without significant loss of accuracy. As the size of the area increases, however, the effects of the earth's curvature become significant and must be taken into account in order to maintain accuracy. Since photogrammetric projects often cover large areas, it is necessary to have quantitative knowledge of the earth's shape in order to determine accurate relative positions of points and to produce reliable maps.

The field of *geodesy* involves the study of the size, shape, gravity, rotation, and crustal movement of the earth. It is a highly refined science which provides the basis for earth-related reference coordinate systems. To understand its basis, three "reference surfaces" must be defined: the *physical earth*, the *geoid*, and the *ellipsoid*. These are illustrated in Fig. 5-1. The physical earth is exactly what the term implies. It contains mountains, valleys, plains, and ocean floors, and is the primary object which is to be mapped. Although generally viewed as a static object, the physical earth is subject to subtle crustal movements, which are accounted for in modern geodesy.

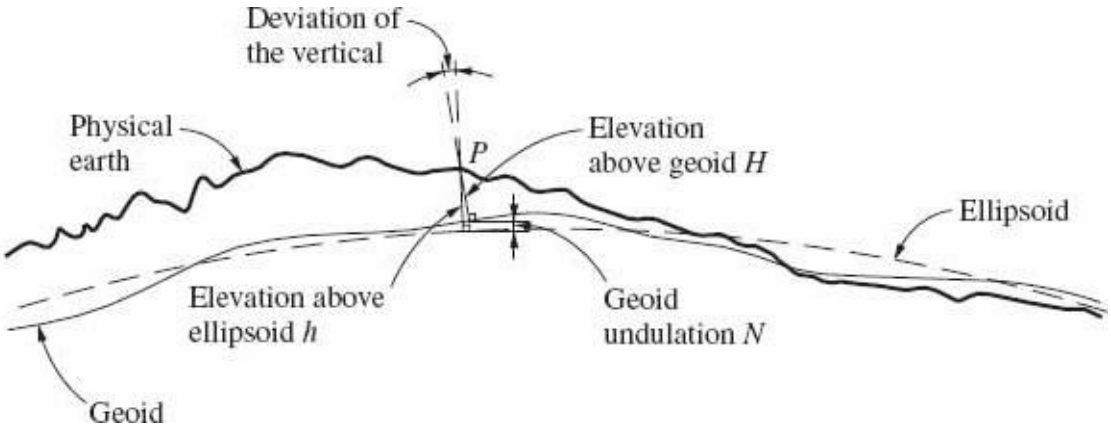


FIGURE 5-1 The three fundamental geodetic reference surfaces: physical earth, geoid, and ellipsoid.

The *geoid* is an equipotential gravity surface, which is generally considered to be *mean sea level*. It can be imagined as the surface that the earth's seas would form if wind, wave, and tidal action ceased and the oceans were connected through the continents by narrow, frictionless canals. The geoid is a gently undulating surface which is everywhere perpendicular to the direction of gravity. These gentle undulations are due to gravity variations caused by the nonhomogeneous mass distribution of the earth. Note that the amplitude of geoid undulations depicted in Fig. 5-1 are greatly exaggerated. At the scale of this figure, true undulations would result in the geoid and ellipsoid appearing to coincide. The shape of the geoid, in fact, results from the *net attraction*, comprised of gravity and the effect of the earth's rotation.

A reference *ellipsoid* is a mathematically defined surface which approximates the geoid either globally or in a large local area such as a continent. This surface is formed by rotating a two-dimensional ellipse (shown in Fig. 5-2*a*) about its minor axis. Rotation of an ellipse in this manner generates a three-dimensional ellipsoid, as shown in Fig. 5-2*b*. This figure shows the (curved) lines which pass through the north and south poles (NP and SP, respectively), known as *meridians*, and (curved) lines which are parallel to the equator, called *parallels*. A meridian is formed by the intersection of the ellipsoid with a plane containing the pole. A parallel, however, is formed by the intersection of the ellipsoid with a plane that is perpendicular to the pole (i.e., parallel to the equator).

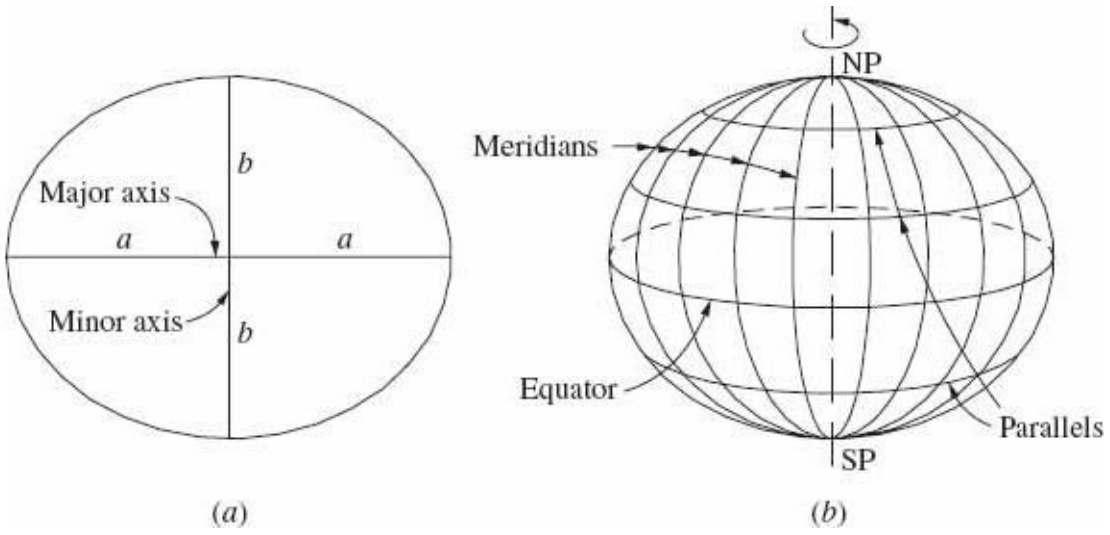


FIGURE 5-2 Definition of a reference ellipsoid. (*a*) Two-dimensional ellipse showing major and minor axes. (*b*) Three-dimensional ellipsoid formed by rotation of ellipse about the minor axis.

To define the size and shape of the reference ellipsoid, at least two constants are required. They are derived from actual measurements made upon the earth. Generally, the semimajor axis of the ellipsoid a and the flattening f are specified. From these two defining constants, other ellipsoid parameters can be derived. The following equations give relationships between the ellipsoid constants of a, the semimajor axis; b, the semiminor axis; f, the flattening; e, the first eccentricity; and e', the second eccentricity.

$$f = 1 - \frac{b}{a} \tag{5-1}$$

$$b = a(1 - f) \tag{5-2}$$

$$e = \frac{\sqrt{a^2 - b^2}}{a} \tag{5-3}$$

$$e^2 = f(2 - f) = \frac{a^2 - b^2}{a^2} \tag{5-4}$$

$$e' = \frac{\sqrt{a^2 - b^2}}{b} \tag{5-5}$$

$$e'^2 = \frac{e^2}{(1 - f)^2} = \frac{a^2 - b^2}{b^2} \tag{5-6}$$

The flattening f is a parameter for an ellipsoid (or ellipse) which quantifies how much it departs from a true sphere (or circle). The value of f for an ellipse can range from 0, which corresponds to a circle, to 1, which corresponds to a completely flattened ellipse, i.e., a straight line. The accepted value of f for the earth is roughly 0.0033, which implies that the earth is very nearly spherical. The first and second eccentricities e and e' are also parameters which quantify how much an ellipse departs from a true circle, with values near 0 denoting near circularity. Table 5-1 gives semimajor axis a and flattening f values for three commonly used reference ellipsoids.

Reference Ellipsoid	Semimajor Axis a	Flattening f
Clarke 1866	6,378,206.4 m	1/294.9786982
GRS80	6,378,137 m	1/298.25722210088
WGS84	6,378,137 m	1/298.257223563

TABLE 5-1 Parameters for Select Reference Ellipsoids

As noted previously, the values for a and f are selected on the basis of geodetic measurements

made in different locations of the world. In the past, reference ellipsoids were derived in order to fit the geoid in certain local regions, such as North America. The Clarke 1866 ellipsoid is an example of this type of local surface, which was a best-fit to the geoid in North America. More recently, given accurate global measurement technology such as GPS, Doppler satellite measurements, and very long baseline interferometry (VLBI), reference ellipsoids have been derived which give a best-fit to the geoid in a worldwide sense. The GRS80 and WGS84 ellipsoids are examples of worldwide reference surfaces.

5-3 Geodetic Coordinate System

Geodetic coordinates for specifying point locations relative to the earth's surface are latitude ϕ, longitude λ, and height h. These coordinates all depend upon a reference ellipsoid for their basis. Latitude and longitude are horizontal components, while the vertical component is height. These three coordinates are illustrated in Fig. 5-3. This figure shows a point P with a line passing through it, perpendicular to the ellipsoid and extending to the polar axis. This line is called the *normal*. The longitude λ of a point is given by the angle in the plane of the equator from the *prime meridian* (usually the meridian through Greenwich, England) to the *local meridian* (meridian passing through the normal line). Values of longitude range from –180° to +180° with those west of the prime meridian being negative and those to the east being positive. The latitude (ϕ) of a point is the angle from the equatorial plane to the normal line. Values of latitude range from –90° to +90° with those north of the equator being positive and those to the south being negative. The latitude is more clearly illustrated in Fig. 5-4, which shows a plane section through the ellipsoid containing the local meridian and the normal.

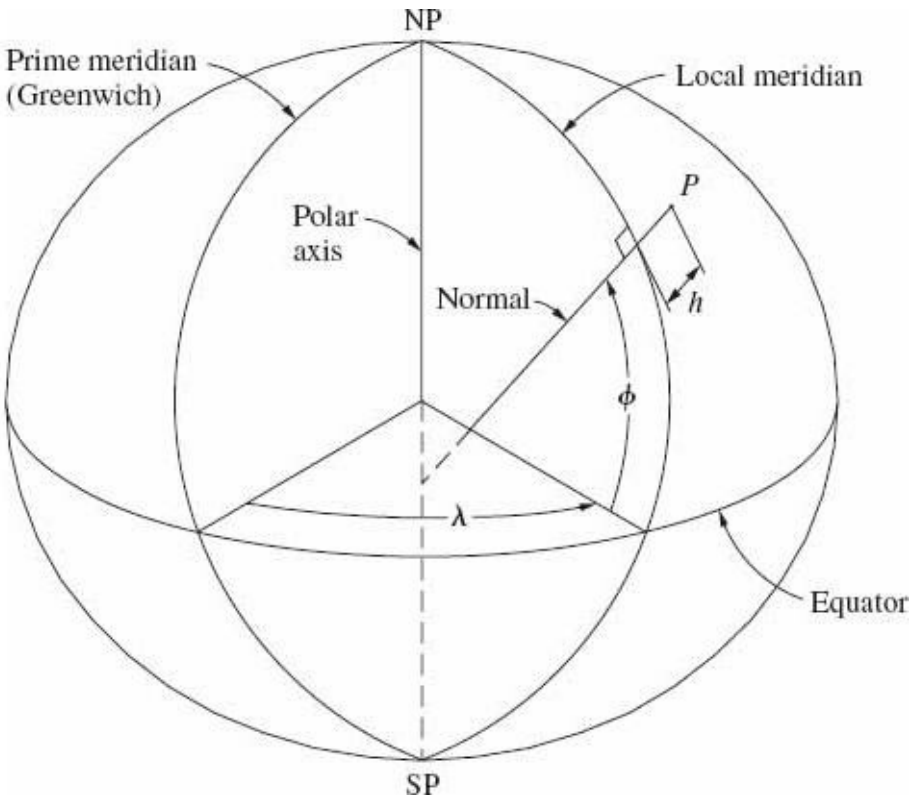


FIGURE 5-3 Geodetic coordinates of latitude ϕ, longitude λ, and height h.

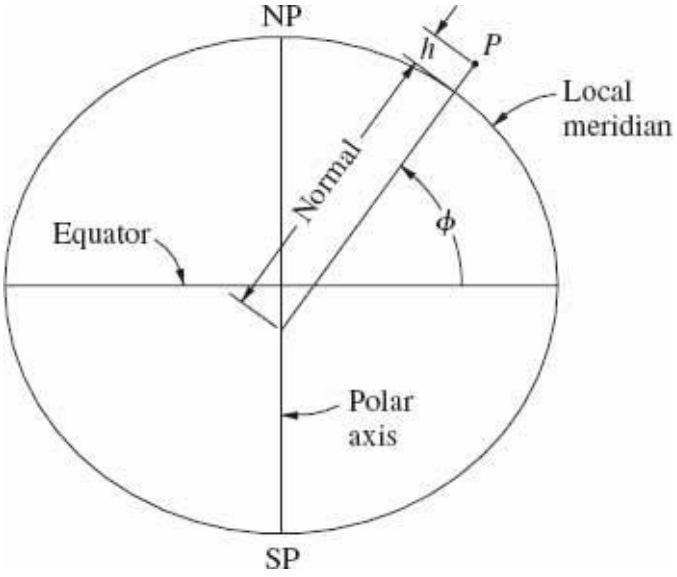


FIGURE 5-4 Illustration of latitude ϕ, normal, and height h in the plane of the local meridian.

As also illustrated in Fig. 5-4, height h is the distance from the surface of the ellipsoid to the point P, in the same direction as the normal. This value specifies the elevation of a point above the ellipsoid, also known as the *ellipsoid height*. The elevation of a point above the geoid H, also known as *orthometric height*, is commonly considered to be the *mean sea level* elevation. Figure 5-1 illustrates the relationship between these two height definitions. The difference between the two heights is referred to as the *geoid undulation* or *geoid height* and is indicated in Fig. 5-1 by the value N. The relationship between ellipsoid height h, orthometric height H, and geoid undulation N is specified in Eq. (5-7).

$$h = H + N \tag{5-7}$$

Figure 5-1 shows two lines through point P perpendicular to the ellipsoid and geoid, respectively. These lines intersect at an angle known as the *deviation of the vertical*. The deviation of the vertical can be determined by precise surveying techniques, and its value never exceeds 2 arc minutes anywhere on the earth.

5-4 Geocentric Coordinates

While geodetic coordinates $\phi\lambda h$ provide an earth-based definition for a point's three-dimensional position, they are related to a curved surface (reference ellipsoid). These coordinates are therefore nonorthogonal and as such are unsuitable for analytical photogrammetry, which assumes a rectangular or cartesian coordinate system. The geocentric coordinate system, on the other hand, is a three-dimensional XYZ cartesian system which provides an earth-centered definition of position, independent of any reference surface. This system has its XY plane in the plane of the equator with the Z axis extending through the north pole. The X axis is oriented such that its positive end passes through the prime meridian. Figure 5-5 illustrates the geocentric coordinate system and its relationship to geodetic coordinates.

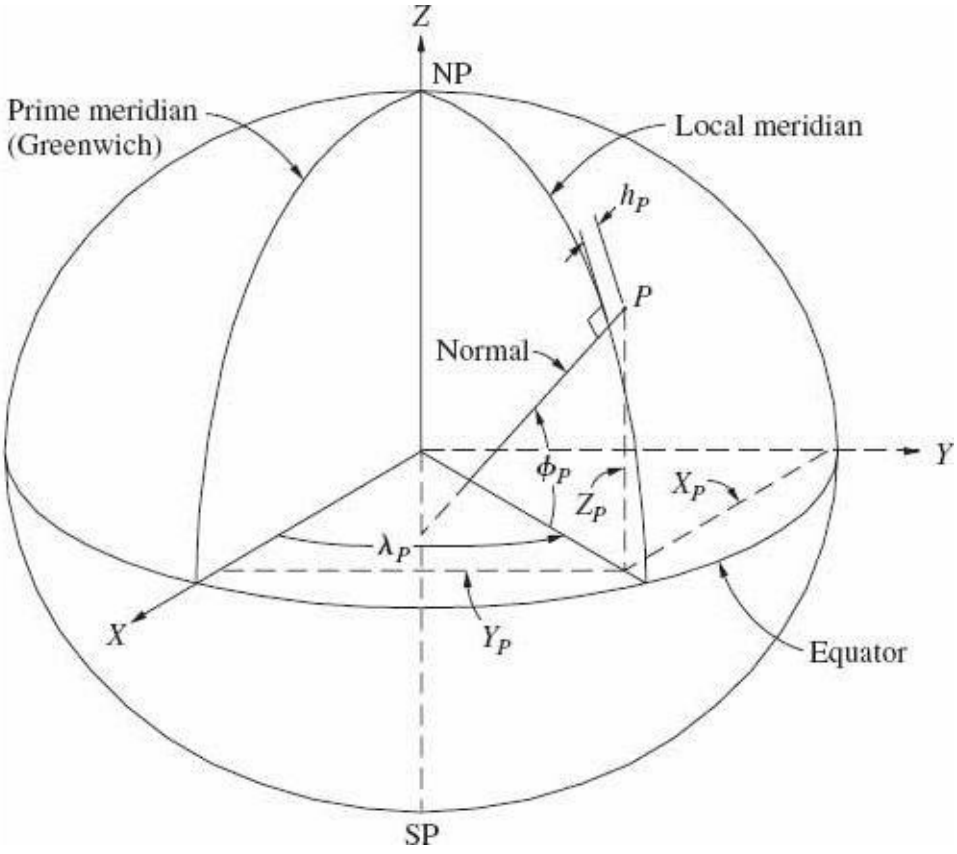


FIGURE 5-5 Relationship between geocentric and geodetic coordinates.

The geocentric coordinate system is a convenient system for many worldwide geodetic applications such as satellite geodesy. Its use in photogrammetry and other mapping applications is not as advantageous since values of the coordinates are very large and have no obvious relationship to the cardinal directions in a local area. Also, the direction of the camera axis would be quantified relative to the earth's pole instead of the local vertical. For these reasons, a special coordinate system, the local vertical system, with its origin in the local project area is generally used.

5-5 Local Vertical Coordinates

A local vertical coordinate system is a three-dimensional cartesian *XYZ* reference system which has its origin placed at a specific point within the project area. At this local origin, the *Z* axis extends straight up from the ellipsoid in the same direction as the normal at the origin. The positive *X* and *Y* axes are tangent to the ellipsoid and point to the east and north, respectively. Figure 5-6 shows the local vertical system and its relationship to geocentric and geodetic coordinates. In this figure, the position of the local origin is specified in terms of geodetic coordinates ϕ_o, λ_o, and h_o, with the last equal to zero. As shown in Fig. 5-6, the local origin has geocentric coordinates X_o, Y_o, and Z_o, and point *P* in the project area has local vertical coordinates $^{X_{lp'}}$ $^{Y_{lp'}}$ and $^{Z_{lp}}$.

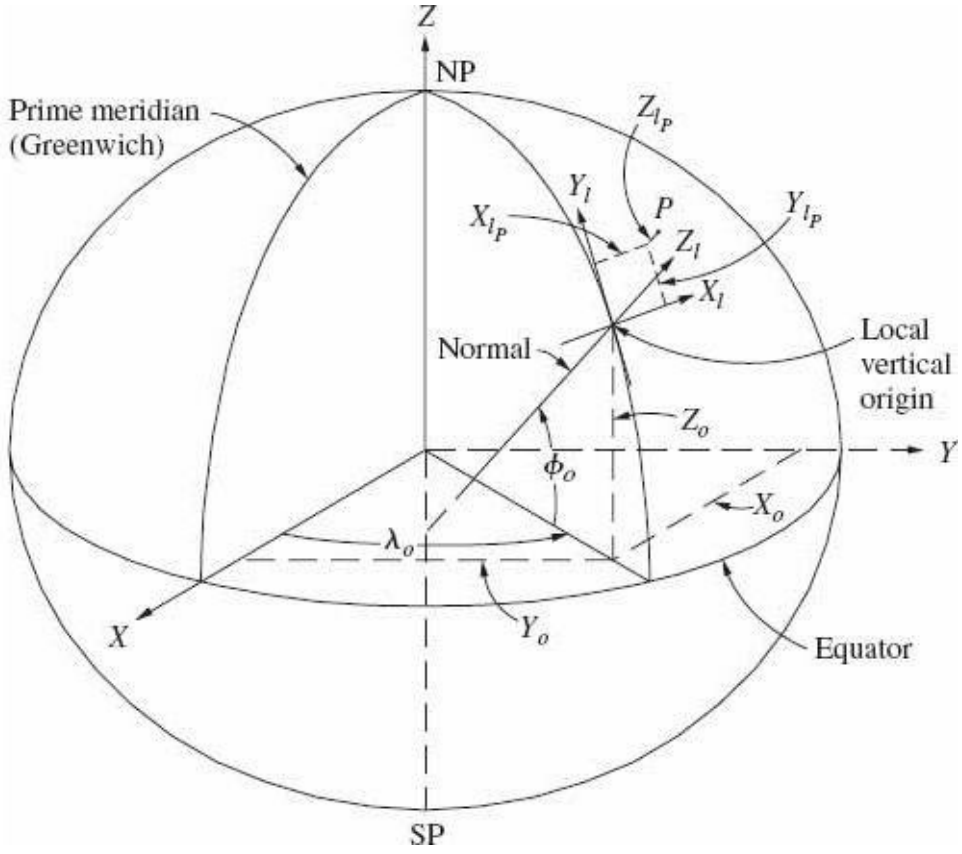


FIGURE 5-6 Local vertical coordinate system relative to geocentric and geodetic systems.

Local vertical coordinates have the characteristic that *X*, *Y*, and *Z* values will correspond roughly to eastings, northings, and heights above the ellipsoid, respectively. However, the farther a point lies from the local origin, the less accurate these correspondences will be. This will have no detrimental effect on coordinates computed through analytical photogrammetry, since the local vertical coordinates can be accurately converted to useful geodetic coordinates or map projection eastings and northings. Appendix F contains detailed descriptions of these conversion procedures, complete with examples.

5-6 Map Projections

One of the fundamental products produced through photogrammetry is a map. A map, in general, consists of points, lines, arcs, symbols, or images which are placed on a flat, two-dimensional surface such as a sheet of paper or computer display. It is, in a manner of speaking, a scaled representation of what a human would see while walking from point to point in the mapped area. As such, it is preferable that the map present the viewpoint from directly overhead throughout the area. In mapping the earth's surface, it is impossible to achieve this overhead viewpoint on a two-dimensional medium without distortion because of the earth's curved shape. Map projections were created to accomplish this viewpoint with a carefully defined and understood amount of distortion.

While there are several types of map projections to choose from, those most often used in photogrammetric mapping applications are *conformal*, meaning that true shape is maintained. Two particular conformal map projections which will be discussed here are the *Lambert conformal conic* and the *transverse Mercator*. These map projections both employ the concept of the *developable surface*. A developable surface is a surface that may be three-dimensional in its natural form, but can be "unrolled" and laid flat. The developable surface is created which nearly coincides with the

ellipsoid in the region being mapped. Points are then projected from the ellipsoid onto the developable surface or vice versa.

As its name implies, the Lambert conformal conic projection uses a cone as its developable surface. The axis of the cone is made to coincide with the minor axis of the ellipsoid and will pass through the ellipsoid along two parallels of latitude, called the *standard parallels*. Figure 5-7 shows a cone superimposed on the reference ellipsoid of the earth. Note that in the region between the standard parallels, the conic surface is below the ellipsoid; therefore, lines that are projected from the ellipsoid to the cone will be made shorter, and those outside the standard parallels will be made longer. This change in dimension can be quantified by a *scale factor* which is less than 1 between the standard parallels, greater than 1 outside, and exactly equal to 1 at the standard parallels. Since this scale factor varies in the north-south direction but remains the same in the east-west direction, the Lambert conformal conic projection is appropriate for areas of limited extent north-south, but wide extent east-west. It is well-suited for an area such as the state of Tennessee.

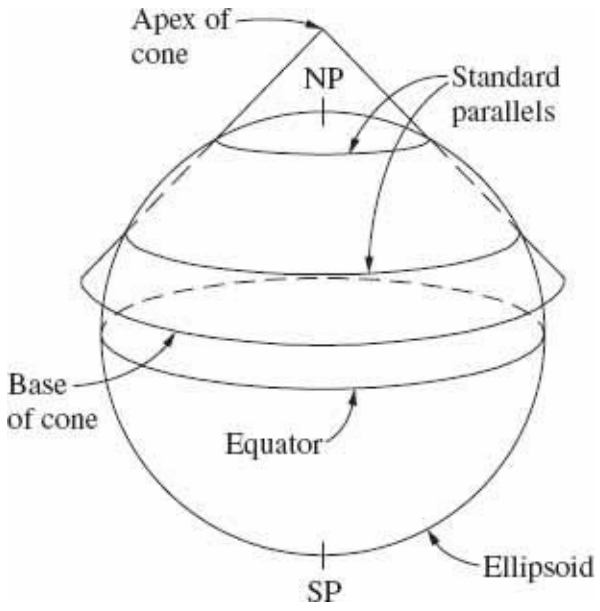


FIGURE 5-7 Cone used in the Lambert conformal conic projection.

The *XY* (easting, northing) coordinates of points within a Lambert conformal conic projection are based on the cone after it has been unrolled and laid flat. Figure 5-8 shows how the flattened cone is placed relative to the *XY* axes. In this figure, parallels of latitude form concentric circular arcs centered on the apex of the cone, and meridians appear as lines which are radial from the apex.

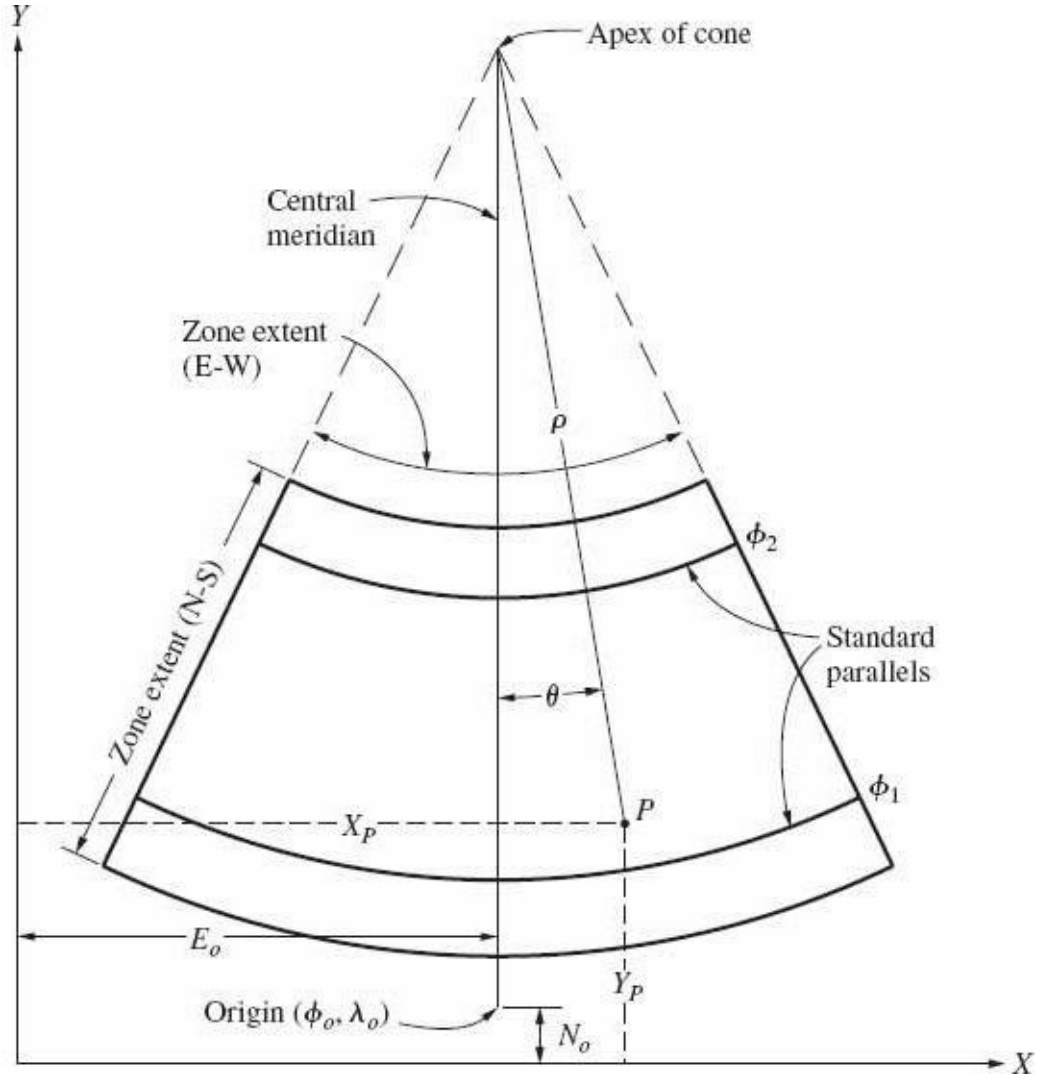


FIGURE 5-8 The Lambert cone unrolled and laid flat.

Different Lambert conformal conic projections can be set up for specific local areas or "zones." A zone is typically defined in terms of the extent of its latitude and longitude. When a Lambert conformal conic map projection is developed for a specific area, a *central meridian* is selected whose longitude is equal to that of the approximate center of the zone. The *origin* for the map projection is also selected. It lies on the central meridian at a location below the coverage of the zone. Figure 5-8 illustrates both the central meridian and the origin. Six parameters uniquely define a Lambert conformal conic map projection for a specific zone. Two of the parameters are the latitudes ϕ_1 and ϕ_2 of the two standard parallels depicted in Figs. 5-7 and 5-8. Two others are the latitude and longitude ϕ_o and λ_o, respectively, of the grid origin. The final two parameters are the *false easting* E_o and *false northing* N_o of the origin. These latter parameters, as shown in Fig. 5-8, are included so that all coordinates within the zone will be positive. A point P having a particular latitude and longitude (ϕ_P, λ_P) will have corresponding map projection coordinates (X_P,Y_P).

Conversion of latitude and longitude ($\phi\lambda$) of a point into map projection coordinates (XY) is known as the *forward* or *direct* conversion. Computation of the forward conversion for the Lambert conformal conic projection involves complex mathematics which are given in Sec. F-5. Conversion from XY to $\phi\lambda$ is referred to as the *inverse* conversion, procedures for which are also given in Sec. F-5.

Another commonly used map projection is the transverse Mercator, which employs a cylinder as its developable surface. For an assumed spherical earth, the developable surface is a true cylinder; however, using an ellipsoidal representation of the earth requires the use of a cylinder that has been

flattened slightly to conform to the shape of the ellipsoid. While a term such as “cylindroid” might be more appropriate for the developable surface, it would be inconsistent with practically all published literature on the subject and thus the word “cylinder” will be used in this text. The axis of the cylinder is defined so as to lie in the plane of the equator, transverse to the minor axis of the ellipsoid. Figure 5-9 shows a cylinder superimposed on the reference ellipsoid of the earth. The cylinder intersects the reference ellipsoid along two rings which are nominally oriented in the north-south direction. Note that in the region between the rings of intersection, the cylinder is below the ellipsoid and therefore lines that are projected from the ellipsoid to the cylinder will be made shorter, and those outside the rings will be made longer. This is similar to the condition with regard to standard parallels of the Lambert conformal conic projection. This change in dimension can also be quantified by a scale factor which is less than 1 between the rings of intersection, greater than 1 outside, and exactly equal to 1 at the intersecting rings. Since this scale factor varies in the east-west direction but remains approximately the same in the north-south direction, the transverse Mercator projection is appropriate for areas of limited extent east-west, but with long extent north-south. It is well suited for areas such as the states of Vermont and New Hampshire.

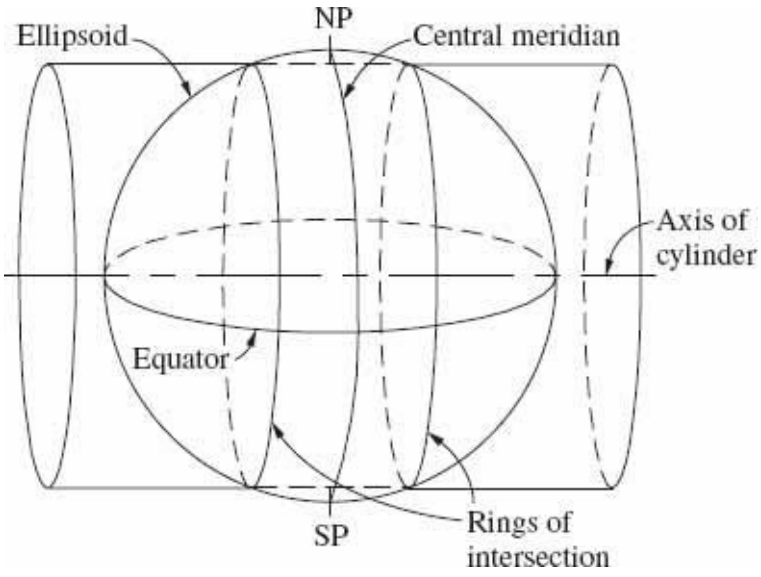


FIGURE 5-9 Cylinder used in the transverse Mercator projection.

The *XY* (easting, northing) coordinates are based on the cylinder after it has been unrolled and laid flat. Figure 5-10 shows how the flattened cylinder is placed relative to the *X* and *Y* axes. In this figure, parallels of latitude and meridians of longitude appear as lines which take the shapes of complex curves, symmetric about the central meridian.

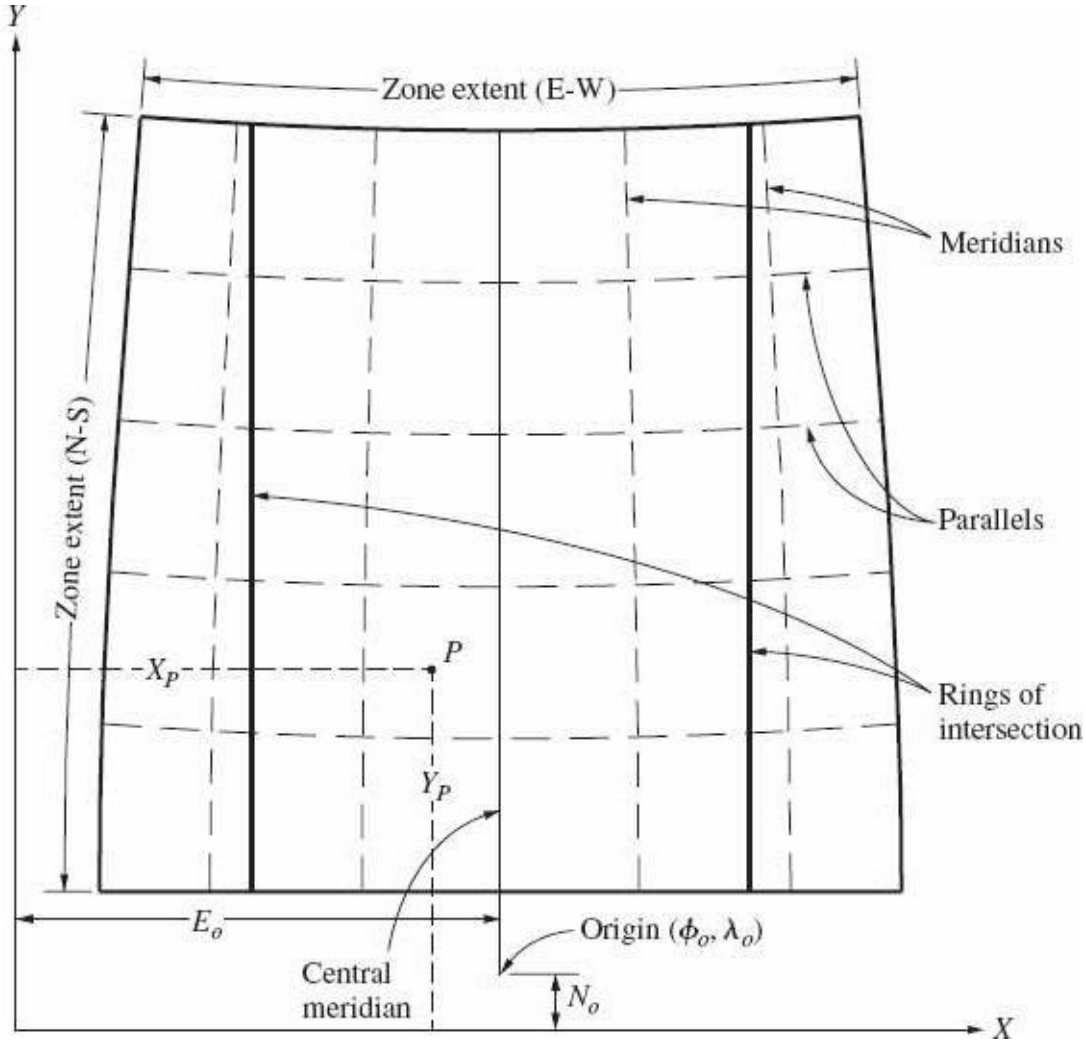


FIGURE 5-10 The transverse Mercator cylinder unrolled and laid flat.

As with the Lambert conformal conic projection, different transverse Mercator projections can be set up in local areas or zones, also defined in terms of their ranges of latitude and longitude. To develop a transverse Mercator map projection for a specific area, a central meridian is selected in the approximate center of the zone. An origin is also defined which lies on the central meridian at a location below the coverage of the zone (see Fig. 5-10). Five parameters uniquely define a transverse Mercator map projection for a specific zone. The first two parameters are the latitude and longitude, ϕ_o and λ_o, of the grid origin. A third parameter, k_o, is the scale factor along the central meridian. Finally, as shown in Fig. 5-10, a false easting E_o and false northing N_o of the origin are included to keep the coordinates positive. A point P having a particular latitude and longitude (ϕ_P, λ_P) will have corresponding map projection coordinates (X_P, Y_P).

The forward conversion of latitude and longitude ($\phi\lambda$) of a point to XY coordinates for the transverse Mercator projection involves mathematical developments which are even more complex than those of the Lambert conformal conic projection (see Sec. F-6). The inverse conversion (from XY to $\phi\lambda$) is equally complex, and these procedures are also given in Sec. F-6.

Both the Lambert conformal conic and transverse Mercator projections are used in *state plane coordinate* (SPC) systems in the United States. These SPC systems were established to provide convenient local coordinate systems for surveying and mapping. In the SPC system, each state is divided into one or more zones chosen so that the maximum scale distortion is no more than 1 part in 10,000. To achieve this distortion limit, the north-south dimension of each Lambert conformal conic zone and the east-west dimension of each transverse Mercator zone are limited to approximately 254

kilometers (km). Each zone has its own unique set of defining parameters, with some states, such as Rhode Island, having a single zone, and other states, such as Alaska, having as many as 10 zones. (Note that one of the Alaska zones uses the oblique Mercator projection.)

Another common map projection system is the *universal transverse Mercator* (UTM) system. This system was established to provide worldwide coverage between 80° S and 80° N latitude by defining 60 zones, each having a 6° longitude range. UTM zone 1 extends from 180° west longitude to 174° west longitude, with a central meridian of 177° west. Zone numbers increase to the east, at an equal spacing of 6° longitude. For example, zone 17 extends from 84° west to 78° west and has a central meridian of 81° longitude. The value of the scale factor along the central meridian k_o is equal to 0.9996 for every zone, resulting in a maximum scale distortion of 1 part in 2500. Each zone has its origin (ϕ_o λ_o) at the intersection of the equator with the central meridian. The false easting for each zone is 500,000 m; for latitudes north of the equator, the false northing is 0 m, and for latitudes south of the equator, the false northing is 10,000,000 m.

The Lambert conformal conic and transverse Mercator map projections are by no means the only map projections available. Other projections which may be routinely encountered are the *polyconic, polar stereographic*, and *space oblique Mercator*. Although discussion of these projections is beyond the scope of this text, the reader is encouraged to consult references at the end of this chapter for details of these and other map projections.

5-7 Horizontal and Vertical Datums

A *datum* is a system of reference for specifying the spatial positions of points. These spatial positions are generally expressed in terms of two separate components, *horizontal* and *vertical*. Thus datums have also traditionally been of two kinds: horizontal and vertical. In a physical sense, a datum consists of a network of well-distributed control monuments whose positions have been determined through precise control surveys. These monuments serve as reference points for originating subordinate surveys of all types. Thus they provide the basis for specifying the relative positions of points both for surveying operations and for mapping purposes.

Common horizontal datums encountered in the United States include the *North American Datum of 1927* (NAD27), the *North American Datum of 1983* (NAD83), the *World Geodetic System of 1984* (WGS84), various statewide *high-accuracy reference networks* (HARNs), and the *International Terrestrial Reference Framework* (ITRF). They provide a means of relating horizontal coordinates derived through surveying and mapping processes to established coordinate reference systems.

Although the theoretical development of horizontal datums is rather complicated, conceptually they can be considered to be based on three primary components as minimum constraints: a reference ellipsoid, an origin, and an angular alignment. The North American Datum of 1927, for example, uses the Clarke 1866 ellipsoid as its reference surface which was a best-fit to the geoid in North America. The origin for NAD27 is a point known as “Meades Ranch,” located in the state of Kansas. At this origin the geoid and ellipsoid, as well as their respective normals, were made to coincide; i.e., deviation of the vertical was assumed to be zero there. Based upon the best positional information available at the time, values of latitude and longitude were assigned and fixed for Meades Ranch, which by definition are exact. The angular alignment was established by fixing an azimuth to a nearby point named “Waldo” and by aligning the semiminor axis of the reference ellipsoid with the rotational axis of the earth. A large network of points, spanning the North American continent, were connected to this pair of points through high-accuracy control surveys. All points in the network were monumented to perpetuate their positions. Based on the latitude and longitude coordinates of Meades

Ranch and the azimuth to Waldo, and the control survey data, coordinates of the other points were computed by a least squares adjustment. The resulting coordinates were then made available to the public so that surveyors and mappers throughout North America could establish coordinates relative to NAD27 by originating their surveys from monumented datum points in the vicinity.

As noted earlier, in a physical sense a horizontal datum such as NAD27 consists of the monument points along with their published coordinates. Since these coordinates were computed from a large number of measurements, each of which contains a certain amount of error, NAD27 likewise contains distortions due to these errors. In addition, subsequent surveys expanded NAD27 by connecting to previously established points in a piecemeal fashion. With advances in instrumentation, particularly accurate electronic distance-measuring devices, the distortions inherent in NAD27 began to cause difficulties in constraining the newer, more accurate measurements to the distorted system. To address this problem, the U.S. National Geodetic Survey created a new datum known as NAD83. This improved datum included virtually all points from NAD27 and its historic measurements, as well as nearly one-quarter million new points created from electronically measured distances and satellite observations. This entire system of combined measurements was adjusted by least squares. The values of computed coordinates were first published in 1986, 3 years later than planned, but nevertheless the system was called NAD83.

The reference ellipsoid of the NAD83 datum is the Geodetic Reference System of 1980 (GRS80) which was a worldwide best-fit to the geoid. The point of origin for NAD83 was made to coincide with the mass center of the earth, which was indirectly determined on the basis of satellite orbits. The semiminor axis of the reference ellipsoid was aligned with the rotational axis of the earth, and an internationally accepted basis for longitude was also used. Since this new datum included many new measurements and had a different set of constraints, points that were common to both datums now had two sets of horizontal coordinates. In some cases they differed by as much as 150 m. For this reason it is essential that the reference datum for any survey or map product be clearly indicated so as to avoid confusion.

The World Geodetic System of 1984 was established by the U.S. Department of Defense during the same time period that NAD83 was being developed. This datum is employed for all maps and charts produced for use by the U.S. armed forces. WGS84 is also the datum to which the broadcast ephemeris of the *Global Positioning System* (GPS) is referenced (see Sec. 16-7).

At the same time that NAD83 was being completed, GPS was beginning to be widely used for geodetic surveys. Due to the exceptionally high accuracy of GPS, discrepancies were being revealed in the newly created NAD83. As use of GPS expanded, these discrepancies became a significant nuisance for geodesists, and newer, more accurate datums were sought. As a response, the National Geodetic Survey, in cooperation with individual states, began to establish high-accuracy reference networks (HARNs). These networks consisted of newly established points which were surveyed using high-accuracy GPS observations. The results were statewide networks of points whose relative positions are known to an accuracy of better than 1 part per million. The HARNs were individually connected to a high-accuracy worldwide network; therefore even though these networks have a high degree of internal consistency, there are discontinuities along the borders between states.

The previously mentioned horizontal datums are static systems; i.e., the coordinates of the monument points are based on a specific moment in time. It has been well established that the surface of the earth is dynamic, subject to crustal movements of several centimeters per year and perhaps more. In addition, the rotational axis of the earth is continually on the move at a slow but detectable rate. With current high-accuracy surveying techniques such as GPS, the dynamic nature of the earth's surface is readily detectable. In response, the International Earth Rotation Service established the *International Terrestrial Reference Frame*, or ITRF. The published values for this sophisticated datum

consist of geocentric coordinates at a specific point in time (*epoch*) along with velocity vectors which can be used to determine the precise locations of the point at a later time. This ultraprecise datum is accurate to 1 part in 50,000,000, which corresponds to a precision of approximately 15 cm between any two points on the earth regardless of their locations! The ITRF is essentially a three-dimensional reference datum of high accuracy which is commonly used as a basis for precise GPS orbit determination. Periodically, WGS84 has been refined so as to closely coincide with the ITRF.

In order to eliminate the need for separate HARNs for individual states, the National Geodetic Survey has readjusted the NAD83 datum using GPS and other measurements connecting the HARN points along with those for the existing NAD83 points. The result is the unified *National Spatial Reference System* (NSRS). Since this datum uses the GRS80 ellipsoid and retains the original location of the earth's mass center as its origin, it is more consistent with earlier readjustments of NAD83. Because of this and the year of its most recent readjustment, the new datum is commonly designated as NAD83(2011). Compared to the location of the earth's mass center in the ITRF, there is a displacement of approximately 2 m to the NAD83(2011) origin. Likewise, since WGS84 is periodically updated to the latest ITRF epoch, there is a displacement of approximately 2 m between WGS84 and NAD83(2011) origins. As noted in Sec. 5-1, a single common reference system should be used in order to produce consistent spatial data products.

A vertical datum is a reference system for giving elevations of points relative to the geoid (i.e., orthometric heights). Two primary vertical datums are currently in use in the United States: the *National Geodetic Vertical Datum of 1929* (NGVD29) and the *North American Vertical Datum of 1988* (NAVD88).

The NGVD29 was established on the basis of tide gauge measurements at 26 stations on the North American coast, together with interconnecting leveling measurements through a vast network of benchmarks across the continent. At the time this datum was established, it was assumed that the geoid coincided with mean sea level as determined at the tide gaging stations. Elevations were therefore constrained to zero at these stations, and all benchmarks were adjusted to conform to this defined reference. Thus the NGVD29 can be considered to be a mean sea level datum.

The NGVD29 evolved in much the same way as the NAD27 in that many additional vertical surveys were connected to the network in local areas. This fact, in addition to distortions in the datum due to measurement errors and constraint to the tide gauging stations, led to a vertical datum which was not sufficiently accurate for modern applications. Also, the increasing use of GPS dictated the use of a vertical datum that more nearly corresponded to the geoid. Thus the NAVD88 was established. It is based on a worldwide gravity model which is, by definition, the geoid. The zero elevation reference was based on Father Point at Rimouski Canada, near the mouth of the St. Lawrence. Unfortunately this elevation was found to differ from the true mean sea level by approximately 0.5 m. Thus the reference surface of the NAVD88 is slightly different from the true geoid. Despite this inconsistency, the NAVD88 is more compatible with worldwide horizontal datums.

It is often necessary to convert (transform) points that have been referenced in one datum to another. Examples would be to transform from NAD27 to NAD83, or from NGVD29 to NAVD88. These transformations have become especially commonplace with the increasing use of geographic information systems. These systems often utilize information from different dates and different sources, and frequently the information is based on different reference coordinate systems. But the information must all be coordinated in a common reference system before being integrated for analysis and use in a GIS. A number of different mathematical procedures have been used for making these conversions. Unless the transformation procedure appropriately accounts for the distortions in the datums, however, errors on the order of several meters can result in the converted positions. To aid in making accurate horizontal datum conversions, the NGS has developed a program called NADCON.

It can convert horizontal datum coordinates between NAD27 and NAD83 to an accuracy of approximately 15 cm with occasional errors as high as 50 cm. A related program called VERTCON, also available from NGS, performs vertical datum conversions between NGVD29 and NAVD88 to an accuracy of approximately 2 cm. Another useful program is GEOID12A which can be used to compute geoid undulation values *N* within the area encompassed by the NAVD88. These programs are available from the NGS on their website.[1]

References

American Society of Photogrammetry: *Manual of Photogrammetry*, 4th ed., Bethesda, MD, 1980, chaps. 8 and 9.

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, chap. 3.

Bomford, G.: *Geodesy*, 4th ed., Clarendon Press, Oxford, 1980.

Boucher, C., and Z. Altamimi: "International Terrestrial Reference Frame," *GPS World*, vol. 7, no. 9, 1996, p. 71.

Colvocoresses, A. P.: "The Gridded Map," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, 1997, p. 377.

Doyle, F. J.: "Map Conversion and the UTM Grid," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, 1997, p. 367.

Featherstone, W., and R. B. Langley: "Coordinates and Datums and Maps! Oh My!" *GPS World*, vol. 8, no. 1, 1997, p. 34.

Ghilani, C. D., and P. R. Wolf: *Elementary Surveying: An Introduction to Geomatics*, 13th ed., Prentice Hall, New Jersey, 2011.

National Geodetic Survey: "Proceedings of the 2010 Federal Geospatial Summit on Improving the National Spatial Reference System" Silver Spring, MD, http://www.ngs.noaa.gov/2010Summit/2010FederalGeospatialSummitProceedings.pdf.

National Imagery and Mapping Agency: "Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems," NIMA Technical Report 8350.2, 3d ed., Bethesda, MD, 1997.

Pursell, D. G. and M. Potterfield: "NAD 83(NSRS 2007) National Readjustment Final Report," NOAA Technical Report NOS NGS 60, National Geodetic Survey, 2008, http://www.ngs.noaa.gov/PUBS_LIB/NSRS2007/NOAATRNOSNGS60.pdf.

Snyder, J. P.: "Map Projections—A Working Manual," U.S. Geological Survey Professional Paper 1395, U.S. Geological Survey, Washington, 1987.

Stern, J. E.: "State Plane Coordinate System of 1983," NOAA Manual NOS NGS 5, National Oceanic and Atmospheric Administration, Rockville, MD, 1989.

Terry, N. G., Jr.: "Field Validation of the UTM Gridded Map," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, 1997, p. 381.

Thompson, M. M.: *Maps For America*, 3d ed., U.S. Geological Survey, Washington, 1987.

Welch, R., and A. Homsey: "Datum Shifts for UTM Coordinates," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, 1997, p. 371.

Problems

5-1. List and briefly describe the three basic reference surfaces in geodesy.

5-2. Using the values of a and f from Table 5-1, compute the values of the semiminor axis b, and the first eccentricity e for the Clarke 1866 ellipsoid. Express your answers to 10 significant figures.

5-3. Repeat Prob. 5-2, except the b and e values for WGS84 should be computed.

5-4. The ellipsoid height h for a point was determined from GPS observation to be +382.67 m. If the geoid undulation N is equal to +9.44 m, what is the orthometric height H for the point?

5-5. Repeat Prob. 5-4, except that the ellipsoid height is –5.62 m and the geoid undulation is –25.83 m.

5-6. Briefly explain the difficulty associated with using the geocentric coordinate system in photogrammetric applications.

5-7. Illustrate and briefly describe how Z values in a local vertical coordinate system vary with distance from the local vertical origin for points having a constant ellipsoid height.

5-8. Briefly describe the variation in scale factor in the north-south direction for a Lambert conformal conic projection.

5-9. List the types of map projections that are used in the U.S. state plane coordinate system.

5-10. Based on the longitude of the area in which you live, determine the number of your local UTM zone.

5-11. Access the applications NADCON, VERTCON, and GEOID12A from the NGS website *http://www.ngs.noaa.gov*. Using the appropriate application, answer the following:

(a) Using the values of latitude and longitude in NAD83 for a point in your local area, determine the corresponding values in NAD27, using NADCON.

(b) Using the NGVD29 elevation of a point in your local area, determine the corresponding elevation in NAVD88, using VERTCON.

(c) Using the same latitude and longitude as in part (*a*), determine the value of the geoid undulation using GEOID12A.

---

[1] The website address of the NGS is http://www.ngs.noaa.gov.

CHAPTER 6
Vertical Photographs

6-1 Geometry of Vertical Photographs

As described in Chap. 1, photographs taken from an aircraft with the optical axis of the camera vertical or as nearly vertical as possible are called *vertical photographs*. If the optical axis is exactly vertical, the resulting photograph is termed *truly vertical*. In this chapter, equations are developed assuming truly vertical photographs. In spite of precautions taken to keep the camera axis vertical, small tilts are invariably present. For photos intended to be vertical, however, tilts are usually less than 1° and rarely exceed 3°. Photographs containing these small unintentional tilts are called *near-vertical* or *tilted photographs*, and for many practical purposes these photos may be analyzed using the relatively simple "truly vertical" equations of this chapter without serious error.

In this chapter, besides assuming truly vertical photographs, other assumptions are that the photo coordinate axis system has its origin at the photographic principal point and that all photo coordinates have been corrected for shrinkage, lens distortion, and atmospheric refraction distortion.

Figure 6-1 illustrates the geometry of a vertical photograph taken from an exposure station L. The *negative*, which is a reversal in both tone and geometry of the object space, is situated a distance equal to the focal length (distance $o'L$ in Fig. 6-1) above the rear nodal point of the camera lens. The *positive* may be obtained by direct emulsion-to-emulsion "contact printing" with the negative. This process produces a reversal of tone and geometry from the negative, and therefore the tone and geometry of the positive are exactly the same as those of the object space. Geometrically the plane of a contact-printed positive is situated a distance equal to the focal length (distance $o'L$ in Fig. 6-1) below the front nodal point of the camera lens. The same is true for an image obtained with a frame-type digital camera. The reversal in geometry from object space to negative is readily seen in Fig. 6-1 by comparing the positions of object points A, B, C, and D with their corresponding negative positions a', b', c', and d'. The correspondence of the geometry of the object space and the positive is also readily apparent. The photographic coordinate axes x and y, as described in Chap. 4, are shown on the positive of Fig. 6-1.

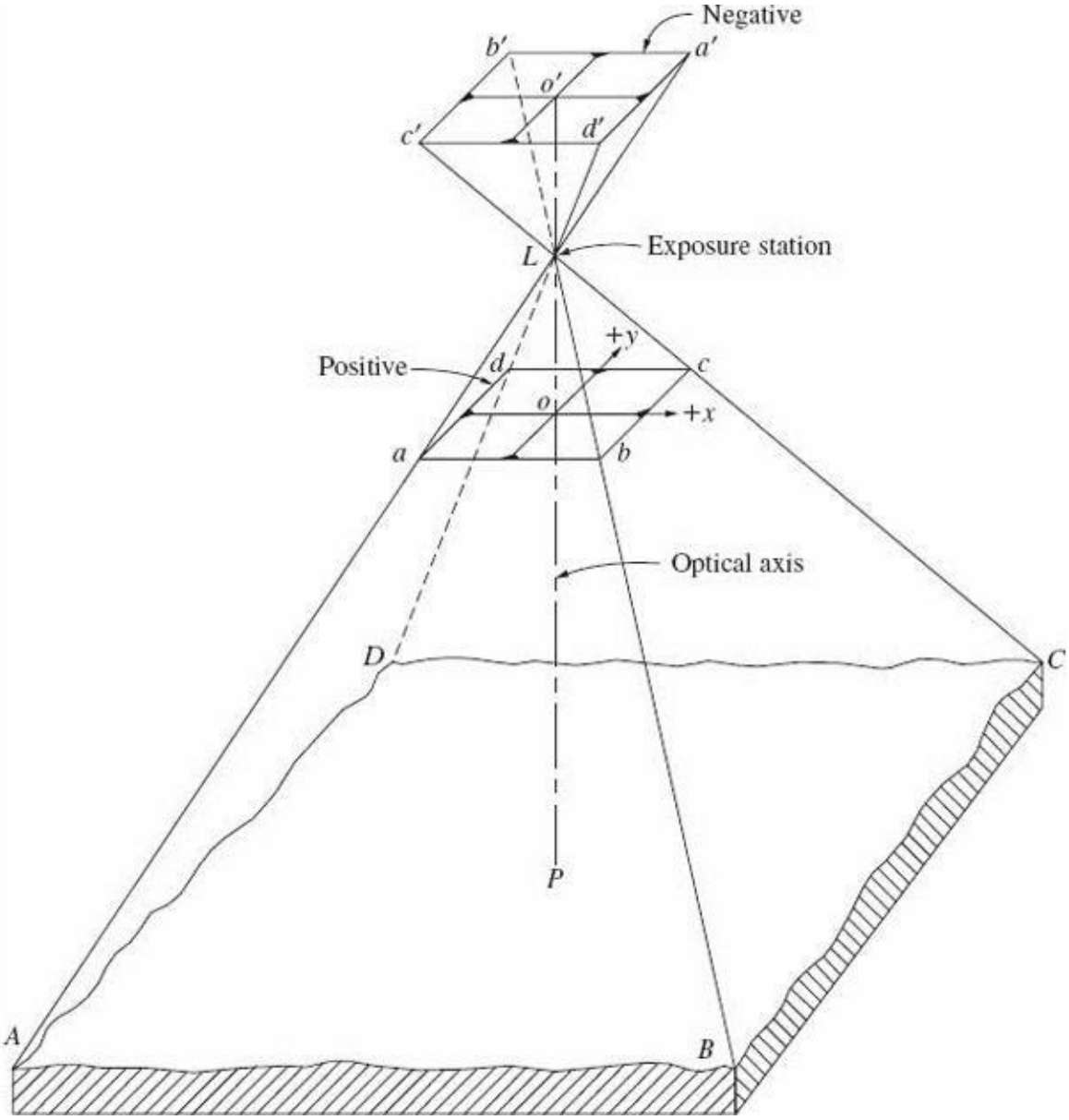


FIGURE 6-1 The geometry of a vertical photograph.

6-2 Scale

Map scale is ordinarily interpreted as the ratio of a map distance to the corresponding distance on the ground. In a similar manner, the *scale of a photograph* is the ratio of a distance on the photo to the corresponding distance on the ground. Due to the nature of map projections, map scale is not influenced by terrain variations. A vertical aerial photograph, however, is a perspective projection, and as will be demonstrated in this chapter, its scale varies with variations in terrain elevation.

Scales may be represented as *unit equivalents, unit fractions, dimensionless representative fractions,* or *dimensionless ratios*. If, for example, 1 inch (in) on a map or photo represents 1000 ft (12,000 in) on the ground, the scale expressed in the aforementioned four ways is

1. Unit equivalents: 1 in = 1000 ft
2. Unit fraction: 1 in/1000 ft
3. Dimensionless representative fraction: 1/12,000
4. Dimensionless ratio: 1:12,000

By convention, the first term in a scale expression is always chosen as 1. It is helpful to remember that

a large number in a scale expression denotes a small scale, and vice versa; for example, 1:1000 is a larger scale than 1:5000.

6-3 Scale of a Vertical Photograph Over Flat Terrain

Figure 6-2 shows the side view of a vertical photograph taken over flat terrain. Since measurements are normally taken from photo positives rather than negatives, the negative has been excluded from this and other figures that follow in this text. The scale of a vertical photograph over flat terrain is simply the ratio of photo distance *ab* to corresponding ground distance *AB*. That scale may be expressed in terms of camera focal length *f* and flying height above ground *H′* by equating similar triangles *Lab* and *LAB* as follows:

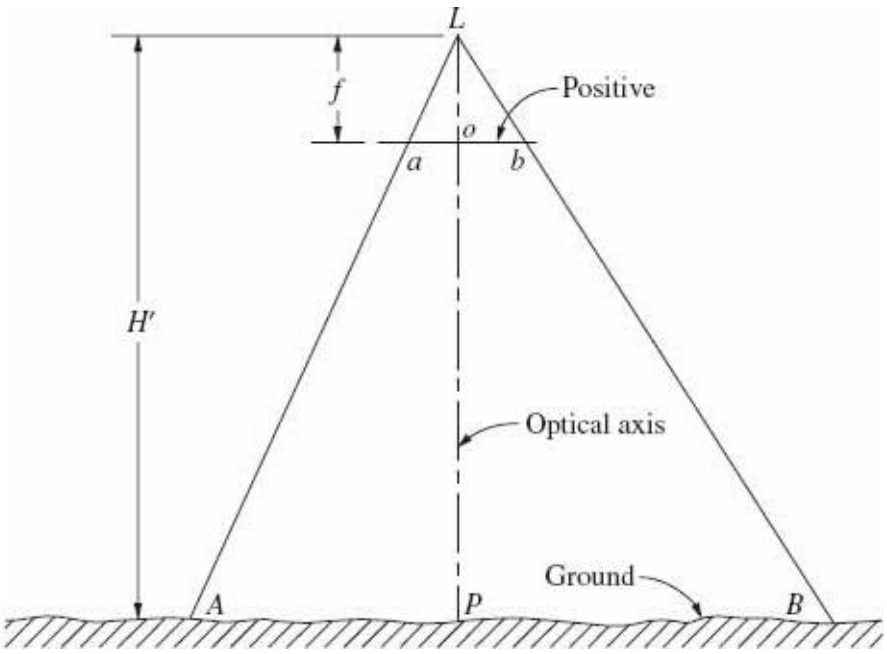


FIGURE 6-2 Two-dimensional view of a vertical photograph taken over flat terrain.

$$S = \frac{ab}{AB} = \frac{f}{H'} \tag{6-1}$$

From Eq. (6-1) it is seen that the scale of a vertical photo is directly proportional to camera focal length (image distance) and inversely proportional to flying height above ground (object distance).

Example 6-1
A vertical aerial photograph is taken over flat terrain with a 152.4 mm-focal-length camera from an altitude of 1830 m above ground. What is the photo scale?

Solution By Eq. (6-1),

$$S = \frac{f}{H'} = \frac{152.4 \text{ mm}}{1830 \text{ m}}$$

$$= \frac{0.1524 \text{ m}}{1830 \text{ m}} = 1/12,\bar{0}00 = 1:12,\bar{0}00 \qquad (1 \text{ in}/10\bar{0}0 \text{ ft})$$

Note the use of the overbar in the solution of Example 6-1 to designate significant figures, as discussed in Sec. A-3. ▲

6-4 Scale of a Vertical Photograph Over Variable Terrain

If the photographed terrain varies in elevation, then the object distance—or the denominator of Eq. (6-1)—will also be variable and the photo scale will likewise vary. For any given vertical photo scale increases with increasing terrain elevation and decreases with decreasing terrain elevation.

Suppose a vertical aerial photograph is taken over variable terrain from exposure station *L* of Fig. 6-3. Ground points *A* and *B* are imaged on the positive at *a* and *b*, respectively. Photographic scale at *h*, the elevation of points *A* and *B*, is equal to the ratio of photo distance *ab* to ground distance *AB*. By similar triangles *Lab* and *LAB*, an expression for photo scale S_{AB} is

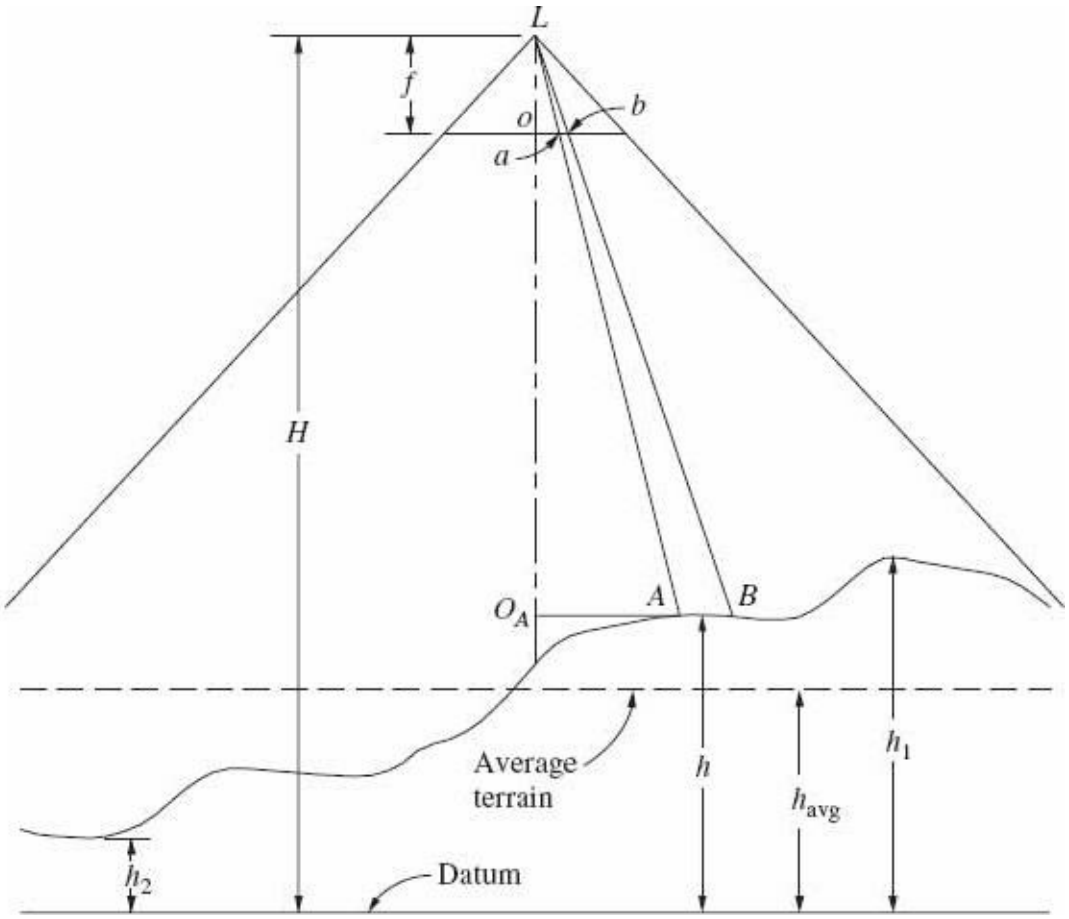


FIGURE 6-3 Scale of a vertical photograph over variable terrain.

$$S_{AB} = \frac{ab}{AB} = \frac{La}{LA} \tag{a}$$

Also, by similar triangles LO_AA and Loa,

$$\frac{La}{LA} = \frac{f}{H-h} \tag{b}$$

Substituting Eq. (*b*) into Eq. (*a*) gives

$$S_{AB} = \frac{ab}{AB} = \frac{La}{LA} = \frac{f}{H-h} \qquad (c)$$

Considering line *AB* to be infinitesimal, we see that Eq. (*c*) reduces to an expression of photo scale at a point. In general, by dropping subscripts, the scale at any point whose elevation above datum is *h* may be expressed as

$$S = \frac{f}{H-h} \qquad (6\text{-}2)$$

In Eq. (6-2), the denominator *H* - *h* is the object distance. In this equation as in Eq. (6-1), scale of a vertical photograph is seen to be simply the ratio of image distance to object distance. The shorter the object distance (the closer the terrain to the camera), the greater the photo scale, and vice versa. *For vertical photographs taken over variable terrain, there are an infinite number of different scales. This is one of the principal differences between a photograph and a map.*

6-5 Average Photo Scale

It is often convenient and desirable to use an *average scale* to define the overall mean scale of a vertical photograph taken over variable terrain. Average scale is the scale at the average elevation of the terrain covered by a particular photograph and is expressed as

$$S_{avg} = \frac{f}{H-h_{avg}} \qquad (6\text{-}3)$$

When an average scale is used, it must be understood that it is exact only at those points that lie at average elevation, and it is an approximate scale for all other areas of the photograph.

Example 6-2

Suppose that highest terrain h_1, average terrain h_{avg}, and lowest terrain h_2 of Fig. 6-3 are 610, 460, and 310 m above mean sea level, respectively. Calculate the maximum scale, minimum scale, and average scale if the flying height above mean sea level is 3000 m and the camera focal length is 152.4 mm.

Solution By Eq. (6-2) (maximum scale occurs at maximum elevation),

$$S_{max} = \frac{f}{H-h_1} = \frac{152.4 \text{ mm}}{(3000-610)\text{m}} = \frac{0.1524 \text{ m}}{2390 \text{ m}}$$

$$= 1/15{,}700 = 1{:}15{,}700 \qquad (1 \text{ in}/1310 \text{ ft})$$

and (minimum scale occurs at minimum elevation)

$$S_{\min} = \frac{f}{H - h_2} = \frac{152.4\ \text{mm}}{(3000 - 310)\text{m}} = \frac{0.1524\ \text{m}}{2690\ \text{m}}$$

$$= 1/17{,}700 = 1{:}17{,}700 \qquad (1\ \text{in}/1470\ \text{ft})$$

By Eq. (6-3)

$$S_{\text{avg}} = \frac{f}{H - h_{\text{avg}}} = \frac{152.4\ \text{mm}}{(3000 - 460)\text{m}} = \frac{0.1524\ \text{m}}{2540\ \text{m}}$$

$$= 1/16{,}700 = 1{:}16{,}700 \qquad (1\ \text{in}/1390\ \text{ft})$$

In each of Eqs. (6-1), (6-2), and (6-3), it is noted that flying height appears in the denominator. Thus, for a camera of a given focal length, if flying height increases, object distance H - h increases and scale decreases. Figures 6-4*a* through *d* illustrate this principle vividly. Each of these vertical photos was exposed using the very same 23-cm format and 152-mm-focal-length camera. The photo of Fig. 6-4*a* had a flying height of 460 m above ground, resulting in an average photo scale of 1:3000. The photos of Fig. 6-4*b*, *c*, and *d* had flying heights above average ground of 910 m, 1830 m, and 3660 m, respectively, producing average photo scales of 1:6000, 1:12,000, and 1:24,000, respectively. ▲

(*a*)

(*b*)

(*c*)

(*d*)

FIGURE 6-4 Four vertical photos taken over Tampa, Florida, illustrating scale variations due to changing flying heights. (*Courtesy Aerial Cartographics of America, Inc.*)

6-6 Other Methods of Determining Scale of Vertical Photographs

In the previous sections, equations have been developed for calculating the scale of vertical aerial photographs in terms of camera focal length, flying height, and terrain elevation. There are, however, other methods of scale determination which do not require knowledge of these values.

A ground distance may be measured in the field between two points whose images appear on the photograph. After the corresponding photo distance is measured, the scale relationship is simply the ratio of the photo distance to the ground distance. The resulting scale is exact only at the elevation of

the ground line, and if the line is along sloping ground, the resulting scale applies at approximately the average elevation of the two endpoints of the line.

Example 6-3
The horizontal distance *AB* between the centers of two street intersections was measured on the ground as 402 m. Corresponding line *ab* appears on a vertical photograph and measures 95.8 mm. What is the photo scale at the average ground elevation of this line?

Solution

$$S = \frac{ab}{AB} = \frac{95.8 \text{ mm}}{402 \text{ m}}$$

$$= \frac{0.0958 \text{ m}}{402 \text{ m}} = \frac{1}{4200} = 1{:}42\bar{0}0 \qquad (1 \text{ in}/35\bar{0} \text{ ft}) \quad \blacktriangle$$

The scale of a vertical aerial photograph may also be determined if a map covering the same area as the photo is available. In this method it is necessary to measure, on the photograph and on the map, the distances between two well-defined points that can be identified on both photo and map. Photographic scale can then be calculated from the following equation:

$$S = \frac{\text{photo distance}}{\text{map distance}} \times \text{map scale} \tag{6-4}$$

Example 6-4
On a vertical photograph the length of an airport runway measures $16\bar{0}$ mm. On a map that is plotted at a scale of 1:24,000, the runway is measured as 103 mm. What is the scale of the photograph at runway elevation?

Solution From Eq. (6-4),

$$S = \frac{160 \text{ mm}}{103 \text{ mm}} \times \frac{1}{24{,}000}$$

$$= 1/15{,}400 = 1{:}15{,}400 \qquad (1 \text{ in}/1290 \text{ ft}) \quad \blacktriangle$$

The scale of a vertical aerial photograph can also be determined without the aid of a measured ground distance or a map if lines whose lengths are known by common knowledge appear on the photo. “Section lines” of a known 1-mile (1-mi) length, or a football field or baseball diamond, could be measured on the photograph, for example, and photographic scale could be calculated as the ratio of the photo distance to the known ground distance.

Example 6-5
What is the scale of a vertical aerial photograph on which a section line measures 151 mm?

Solution The length of a section line is assumed to be 5280 ft. (Actually it can vary considerably from that value.) Photo scale is simply the ratio of the measured photo distance to the ground distance, or

$$S = \frac{151 \text{ mm}}{5280 \text{ ft}} \times \frac{1 \text{ ft}}{304.8 \text{ mm}} = 1{:}10{,}700 \qquad (1 \text{ in}/888 \text{ ft}) \quad \blacktriangle$$

In each of the methods of scale determination discussed in this section, it must be remembered that the calculated scale applies only at the elevation of the ground line used to determine that scale.

6-7 Ground Coordinates from a Vertical Photograph

The ground coordinates of points whose images appear in a vertical photograph can be determined with respect to an arbitrary ground coordinate system. The arbitrary X and Y ground axes are in the same vertical planes as the photographic x and y axes, respectively, and the origin of the system is at the datum principal point (point in the datum plane vertically beneath the exposure station).

Figure 6-5 shows a vertical photograph taken at a flying height H above datum. Images a and b of the ground points A and B appear on the photograph, and their measured photographic coordinates are x_a, y_a, x_b, and y_b. The arbitrary ground coordinate axis system is X and Y, and the coordinates of points A and B in that system are X_A, Y_A, X_B, and Y_B. From similar triangles $La'o$ and $LA'A_o$, the following equation may be written:

$$\frac{oa'}{A_oA'} = \frac{f}{H - h_A} = \frac{x_a}{X_A}$$

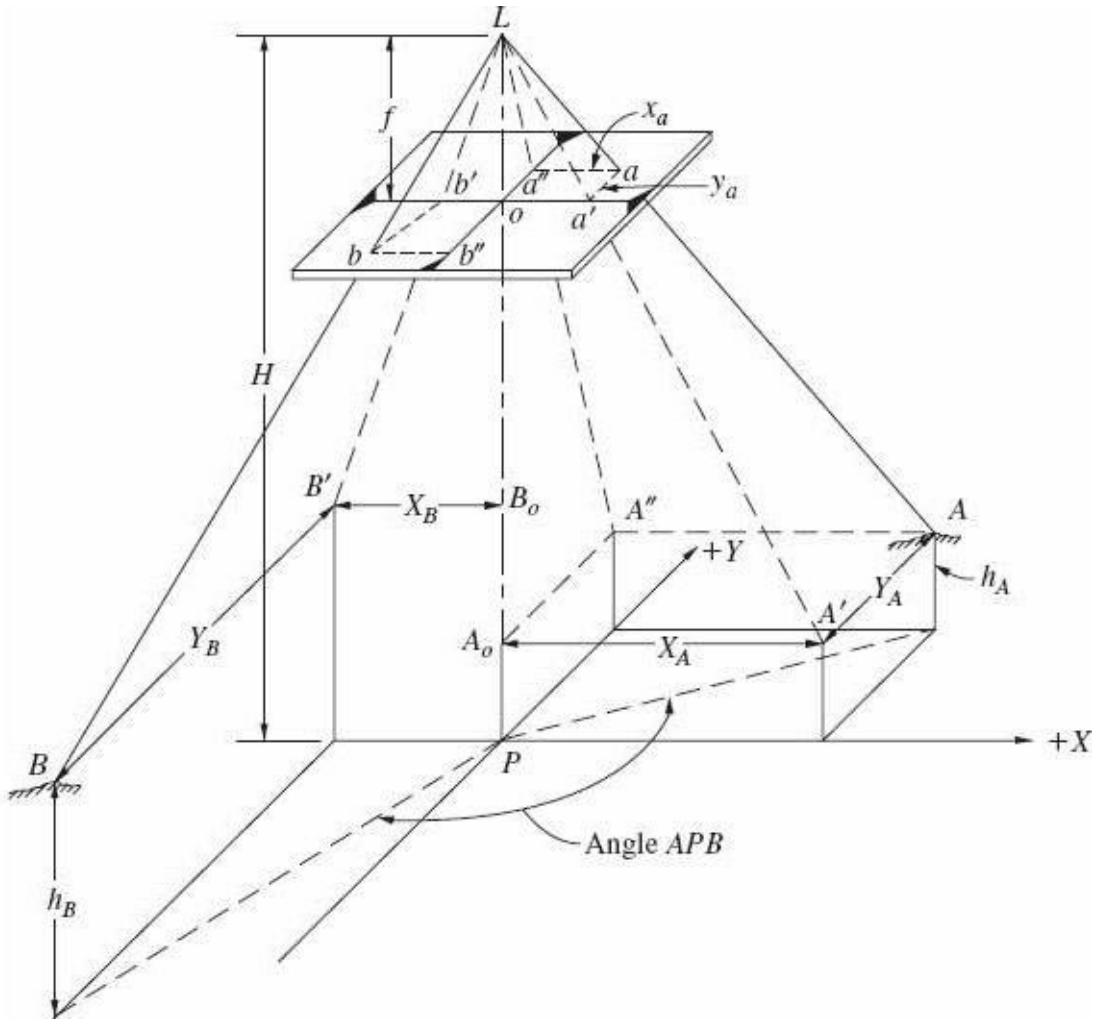


FIGURE 6-5 Ground coordinates from a vertical photograph.

from which

$$X_A = x_a\left(\frac{H-h_A}{f}\right) \tag{6-5}$$

Also, from similar triangles *La″o* and *LA″A*$_o$,

$$\frac{oa''}{A_oA''} = \frac{f}{H-h_A} = \frac{y_a}{Y_A}$$

from which

$$Y_A = y_a\left(\frac{H-h_A}{f}\right) \tag{6-6}$$

Similarly, the ground coordinates of point *B* are

$$X_B = x_b\left(\frac{H-h_B}{f}\right) \tag{6-7}$$

$$Y_B = y_b\left(\frac{H-h_B}{f}\right) \tag{6-8}$$

Upon examination of Eqs. (6-5) through (6-8), it is seen that *X* and *Y* ground coordinates of any point are obtained by simply multiplying *x* and *y* photo coordinates by the inverse of photo scale at that point. From the ground coordinates of the two points *A* and *B*, the horizontal length of line *AB* can be calculated, using the pythagorean theorem, as

$$AB = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2} \tag{6-9}$$

Also, horizontal angle *APB* may be calculated as

$$APB = 90^\circ + \tan^{-1}\left(\frac{X_B}{Y_B}\right) + \tan^{-1}\left(\frac{Y_A}{X_A}\right) \tag{6-10}$$

To solve Eqs. (6-5) through (6-8) it is necessary to know the camera focal length, flying height above datum, elevations of the points above datum, and photo coordinates of the points. The photo coordinates are readily measured, camera focal length is commonly known from camera calibration,

and flying height above datum is calculated by methods described in Sec. 6-9. Elevations of points may be obtained directly by field measurements, or they may be taken from available topographic maps.

Example 6-6
A vertical aerial photograph was taken with a 152.4-mm-focal-length camera from a flying height of 1385 m above datum. Images *a* and *b* of two ground points *A* and *B* appear on the photograph, and their measured photo coordinates (corrected for shrinkage and distortions) are $x_a = -52.35$ mm, $y_a = -48.27$ mm, $x_b = 40.64$ mm, and $y_b = 43.88$ mm. Determine the horizontal length of line *AB* if the elevations of points *A* and *B* are 204 and 148 m above datum, respectively.

Solution From Eqs. (6-5) through (6-8),

$$X_A = \frac{-52.35}{152.4}(1385-204) = -405.7 \text{ m}$$

$$Y_A = \frac{-48.27}{152.4}(1385-204) = -374.1 \text{ m}$$

$$X_B = \frac{40.64}{152.4}(1385-148) = 329.9 \text{ m}$$

$$Y_B = \frac{43.88}{152.4}(1385-148) = 356.2 \text{ m}$$

From Eq. (6-9),

$$AB = \sqrt{(329.9+405.7)^2 + (356.2+374.1)^2} = 1036 \text{ m} \quad \blacktriangle$$

Ground coordinates calculated by Eqs. (6-5) through (6-8) are in an abritrary rectangular coordinate system, as previously described. If arbitrary coordinates are calculated for two or more "control" points (points whose coordinates are also known in an absolute ground coordinate system such as the state plane coordinate system), then the arbitrary coordinates of all other points for that photograph can be transformed to the ground system. The method of transformation used here is discussed in Secs. C-2 through C-5 (of App. C), and an example is given. Using Eqs. (6-5) through (6-8), an entire planimetric survey of the area covered by a vertical photograph can be made.

6-8 Relief Displacement on a Vertical Photograph

Relief displacement is the shift or displacement in the photographic position of an image caused by the relief of the object, i.e., its elevation above or below a selected datum. With respect to a datum, relief displacement is outward for points whose elevations are above datum and inward for points whose elevations are below datum.

The concept of relief displacement is illustrated in Fig. 6-6, which represents a vertical photograph taken from flying height *H* above datum. Camera focal length is *f*, and *o* is the principal point. The image of terrain point *A*, which has an elevation h_A above datum, is located at *a* on the photograph. An imaginary point *A′* is located vertically beneath *A* in the datum plane, and its corresponding imaginary image position is at *a′*. On the figure, both *A′A* and *PL* are vertical lines, and therefore *A′AaLoP* is a vertical plane. Plane *A′a′LoP* is also a vertical plane which is coincident with

A'AaLoP. Since these planes intersect the photo plane along lines *oa* and *oa'*, respectively, line *aa'* (relief displacement of point *A* due to its elevation h_A) is radial from the principal point.

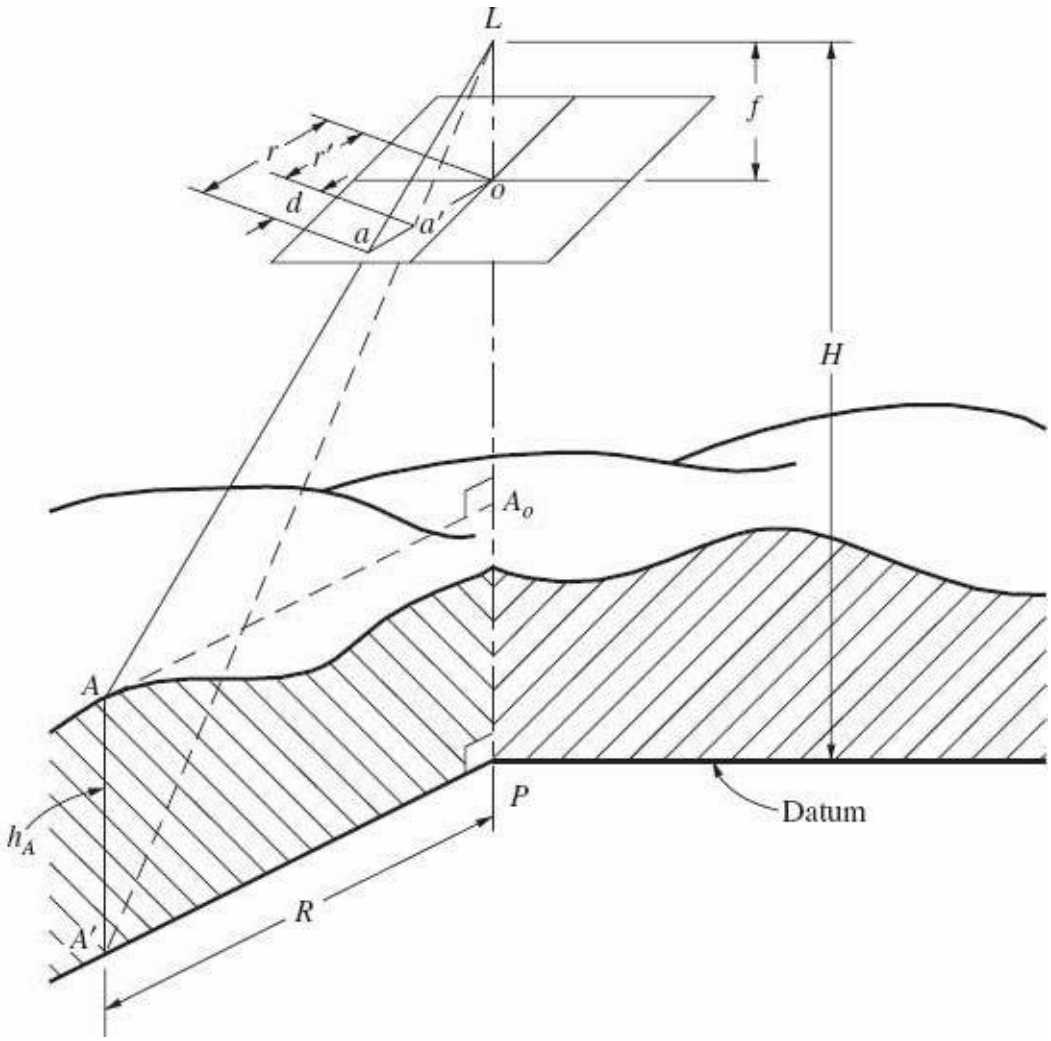


FIGURE 6-6 Relief displacement on a vertical photograph.

An equation for evaluating relief displacement may be obtained by relating similar triangles. First consider planes *Lao* and LAA_o in Fig. 6-6:

$$\frac{r}{R} = \frac{f}{H - h_A} \qquad \text{Rearranging gives} \qquad r(H - h_A) = fR \tag{d}$$

Also, from similar triangles *La'o* and *LA'P*,

$$\frac{r'}{R} = \frac{f}{H} \qquad \text{or} \qquad r'H = fR \tag{e}$$

Equating expressions (*d*) and (*e*) yields

$$r(H - h_A) = r'H$$

Rearranging the above equation, dropping subscripts, and substituting the symbol *d* for *r* - *r'* gives

$$d = \frac{rh}{H} \tag{6-11}$$

where d = relief displacement

h = height above datum of object point whose image is displaced

r = radial distance on photograph from principal point to displaced image (The units of d and r must be the same.)

H = flying height above same datum selected for measurement of h

Equation (6-11) is the basic relief displacement equation for vertical photos. Examination of this equation shows that relief displacement increases with increasing radial distance to the image, and it also increases with increased elevation of the object point above datum. On the other hand, relief displacement decreases with increased flying height above datum. It has also been shown that relief displacement occurs radially from the principal point.

Figure 6-7 is a vertical aerial photograph which vividly illustrates relief displacement. Note in particular the striking effect of relief displacement on the tall buildings in the upper portion of the photo. Notice also that the relief displacement occurs radially from the center of the photograph (principal point). This radial pattern is also readily apparent for the relief displacement of all the other vertical buildings in the photo. The building in the center is one of the tallest imaged on the photo (as evidenced by the length of its shadow); however, its relief displacement is essentially zero due to its proximity to the principal point.

FIGURE 6-7 Vertical photograph of Tampa, Florida, illustrating relief displacements. (*Courtesy US Imaging, Inc.*)

Relief displacement often causes straight roads, fence lines, etc., on rolling ground to appear crooked on a vertical photograph. This is especially true when such roads, fences, etc., occur near the edges of the photo. The severity of the crookedness will depend on the amount of terrain variation. Relief displacement causes some imagery to be obscured from view. Several examples of this are seen in Fig. 6-7; e.g., the street in the upper portion of the photo is obscured by relief displacement of several tall buildings adjacent to it.

Vertical heights of objects such as buildings, poles, etc., appearing on aerial photographs can be calculated from relief displacements. For this purpose, Eq. (6-11) is rearranged as follows:

$$h = \frac{dH}{r} \tag{6-12}$$

To use Eq. (6-12) for height determination, it is necessary that the images of both the top and bottom of the vertical object be visible on the photograph, so that d can be measured. Datum is arbitrarily selected at the base of the vertical object. Consider for example, that the cross-hatched terrain in Fig. 6-6 does not exist and that the surface labeled as datum is the actual terrain. Then line AA' might

represent a tall vertical pole whose top and base are imaged in the photograph. In that case image distances *d* and *r* can be measured and if *H* is known, Eq. (6-12) can be used to determine the height of the pole. Equation (6-12) is of particular import to the photo interpreter, who is often interested in relative heights of objects rather than absolute elevations.

Example 6-7
A vertical photograph taken from an elevation of 535 m above mean sea level (MSL) contains the image of a tall vertical radio tower. The elevation at the base of the tower is 259 m above MSL. The relief displacement *d* of the tower was measured as 54.1 mm, and the radial distance to the top of the tower from the photo center was 121.7 mm. What is the height of the tower?

Solution Select datum at the base of the tower. Then flying height above datum is

$$H = 535 - 259 = 276 \text{ m}$$

By Eq. (6-12),

$$h = \frac{54.1(276)}{121.7} = 123 \text{ m} \quad \blacktriangle$$

6-9 Flying Height of a Vertical Photograph

From previous discussion it is apparent that flying height above datum is an important quantity which is often needed for solving basic photogrammetric equations. Note, for example, that this parameter appears in scale, ground coordinate, and relief displacement equations. For rough computations, flying height may be taken from altimeter readings. Flying heights also may be obtained by using either Eq. (6-1) or Eq. (6-2) if a ground line of known length appears on the photograph. This procedure yields accurate flying heights for truly vertical photographs if the endpoints of the ground line lie at equal elevations. In general, the greater the difference in elevation of the endpoints, the greater the error in the computed flying height; therefore the ground line should lie on fairly level terrain. Accurate results can be obtained by this method, however, even though the endpoints of the ground line are at different elevations, if the images of the endpoints are approximately equidistant from the principal point of the photograph and on a line through the principal point.

Example 6-8
A vertical photograph contains the images of two fence line corners which identify adjacent section corners. The section line lies on fairly level terrain. Find the approximate flying height above the terrain if the camera focal length is 88.9 mm and the section line scales 94.0 mm on the photograph.

Solution Assume the section line to be 5280 ft long. By Eq. (6-1),

$$\frac{94.0 \text{ mm}}{5280 \text{ ft}}\left(\frac{1 \text{ ft}}{304.8 \text{ mm}}\right) = \frac{88.9 \text{ mm}}{H'}\left(\frac{1 \text{ m}}{1000 \text{ mm}}\right)$$

from which

$$H' = 1520 \text{ m above terrain} \quad \blacktriangle$$

Accurate flying heights can be determined even though the endpoints of the ground line lie at different elevations, regardless of the locations of the endpoints in the photo. This procedure requires knowledge of the elevations of the endpoints of the line as well as of the length of the line. Suppose ground line *AB* has its endpoints imaged at *a* and *b* on a vertical photograph. Length *AB* of the ground line may be expressed in terms of ground coordinates, by the pythagorean theorem, as follows:

$$(AB)^2 = (X_B - X_A)^2 + (Y_B - Y_A)^2$$

Substituting Eqs. (6-5) through (6-8) into the previous equation gives

$$(AB)^2 = \left[\frac{x_b}{f}(H - h_B) - \frac{x_a}{f}(H - h_A)\right]^2 + \left[\frac{y_b}{f}(H - h_B) - \frac{y_a}{f}(H - h_A)\right]^2 \tag{6-13}$$

The only unknown in Eq. (6-13) is the flying height *H*. When all known values are inserted into the equation, it reduces to the quadratic form of $aH^2 + bH + c = 0$. The direct solution for *H* in the quadratic is

$$H = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{6-14}$$

Example 6-9

A vertical photograph was taken with a camera having a focal length of 152.3 mm. Ground points *A* and *B* have elevations 437.4 m and 445.3 m above sea level, respectively, and the horizontal length of line *AB* is 584.9 m. The images of *A* and *B* appear at *a* and *b*, and their measured photo coordinates are x_a = 18.21 mm, y_a = –61.32 mm, x_b = 109.65 mm, and y_b = –21.21 mm. Calculate the flying height of the photograph above sea level.

Solution By Eq. (6-13),

$$(584.9)^2 = \left[\frac{109.65}{152.3}(H - 445.3) - \frac{18.21}{152.3}(H - 437.4)\right]^2$$
$$+ \left[\frac{-21.21}{152.3}(H - 445.3) + \frac{61.32}{152.3}(H - 437.4)\right]^2$$

Reducing gives

$$(584.9)^2 = (0.6004H - 268.3)^2 + (0.2634H - 114.1)^2$$

Squaring terms and arranging in quadratic form yields

$$0.4298H^2 - 382.3H - 257{,}100 = 0$$

Solving for *H* by Eq. (6-14) gives

$$H = \frac{382.3 \pm \sqrt{(-382.3)^2 - 4(0.4298)(-257{,}100)}}{2(0.4298)}$$

$$= \frac{382.3 \pm 766.9}{2(0.4298)} = 1337 \text{ m}$$

Note: The positive root was selected, since the negative root yields a ridiculous answer. ▲

6-10 Error Evaluation

Answers obtained in solving the various equations presented in this chapter will inevitably contain errors. It is important to have an awareness of the presence of these errors and to be able to assess their approximate magnitudes. Errors in computed answers are caused partly by random errors in measured quantities that are used in computation, and partly by the failure of certain assumptions to be met. Some of the more significant sources of errors in calculated values using the equations of this chapter are

1. Errors in photographic measurements, e.g., line lengths or photo coordinates
2. Errors in ground control
3. Shrinkage and expansion of film and paper
4. Tilted photographs where vertical photographs were assumed

Sources 1 and 2 can be minimized if precise, properly calibrated equipment and suitable caution are used in making the measurements. Source 3 can be practically eliminated by making corrections as described in Sec. 4-9. Magnitudes of error introduced by source 4 depend upon the severity of the tilt. Generally if the photos were intended to be vertical and if paper prints are being used, these errors are compatible with the other sources. If the photo is severely tilted, or if the highest accuracy is desired, analytical methods of Chap. 11 should be used. For the methods described in this chapter, errors caused by lens distortions and atmospheric refraction are relatively small and can generally be ignored.

A simple and straightforward approach to calculating the combined effect of several random errors is to use statistical error propagation, as discussed in Sec. A-4. This approach involves calculating rates of change with respect to each variable containing error and requires the use of differential calculus. As an example of this approach, assume that a vertical photograph was taken with a camera having a focal length of 152.4 mm. Assume also that a ground distance *AB* on flat terrain has a length of 1524 m and that its corresponding photo distance *ab* measures 127.0 mm. Flying height above ground may be calculated, using Eq. (6-1), as follows:

$$H' = f\left(\frac{AB}{ab}\right) = 152.4\left(\frac{1524}{127.0}\right) = 1829 \text{ m}$$

Now it is required to calculate the expected error dH' caused by errors in measured quantities *AB* and *ab*. This is done by taking partial derivatives with respect to each of these quantities containing error. Suppose that the error σ_{AB} in the ground distance is ±0.50 m and that the error σ_{ab} in the measured photo distance is ±0.20 mm. The rate of change of error in H' caused by the error in the ground length can be evaluated by taking the partial derivative $\partial H'/\partial AB$ as

$$\frac{\partial H'}{\partial AB} = \frac{f}{ab} = \frac{152.4 \text{ mm}}{127.0 \text{ mm}} = 1.200$$

In a similar manner, the rate of change of error in H' caused by the error in the measured image length can be evaluated by taking the partial derivative $\partial H'/\partial ab$ as

$$\frac{\partial H'}{\partial ab} = \frac{-f(AB)}{(ab)^2} = \frac{-152.4 \text{ mm } (1524 \text{ m})}{(127.0 \text{ mm})^2} = -14.40 \text{ m/mm}$$

A useful interpretation of these derivative terms is that an error of 1 m in ground distance *AB* will cause an error of approximately 1.2 m in the flying height, whereas an error of 1 mm in image distance *ab* will cause an error of approximately 14 m in the flying height. Substitution of these derivative terms into the error propagation Eq. (A-2) along with the error terms σ_{AB} and σ_{ab} gives

$$\sigma_{H'} = \sqrt{(1.200)^2(0.50 \text{ m})^2 + (-14.40 \text{ m/mm})^2(0.20 \text{ mm})^2}$$

$$= \sqrt{0.36 \text{ m}^2 + 8.29 \text{ m}^2} = \pm 2.9 \text{ m}$$

Note that the error in H' caused by the error in the measurement of photo distance *ab* is the more severe of the two contributing sources. Therefore, to increase the accuracy of the computed value of H', it would be more beneficial to refine the measured photo distance to a more accurate value. Errors in computed answers using any of the equations presented in this chapter can be analyzed in the manner described above, and the method is valid as long as the contributing errors are small.

References

American Society of Photogrammetry: *Manual of Photogrammetry*, 3d ed., Bethesda, MD, 1966, chap. 2.

———: *Manual of Photogrammetry*, 4th ed., Bethesda, MD, 1980, chap. 2.

Ghilani C.: *Adjustment Computations: Spatial Data Analysis*, Wiley and Sons, Hoboken, NJ, 2010.

Problems

Express answers for scale as dimensionless ratios and answers for distance in meters, unless otherwise specified.

6-1. The photo distance between two image points *a* and *b* on a vertical photograph is *ab*, and the corresponding ground distance is *AB*. What is the photographic scale at the elevation of the ground line?

(*a*) *ab* = 3.57 in; *AB* = 779 ft
(*b*) *ab* = 1.75 in; *AB* = 626 ft
(*c*) *ab* = 6.12 in; *AB* = 5847 ft
(*d*) *ab* = 4.48 in; *AB* = 9287 ft

6-2. Repeat Prob. 6-1, using the values of *ab* and *AB* as indicated.

(*a*) *ab* = 189.5 mm; *AB* = 501.3 m
(*b*) *ab* = 148.3 mm; *AB* = 1043 m

(*c*) *ab* = 195.7 mm; *AB* = 1424 m
(*d*) *ab* = 130.2 mm; *AB* = 1.00 mi

6-3. On a vertical photograph a section line measures 80.3 mm. What is the photographic scale at the elevation of the section line?

6-4. On a vertical photograph a college football field measures 39.0 mm from goal line to goal line (100.0 yd). What is the scale of the photograph at the elevation of the football field?

6-5. A semitractor and trailer combination which is known to be 18 m long measures 11.3 mm long on a vertical aerial photo. What is the scale of the photo at the elevation of the semi?

6-6. Repeat Prob. 6-5, except that a railroad boxcar of 25.0-m known length measures 10.4 mm on the photo.

6-7. An interstate highway pavement of known 24.0-ft width measures 3.57 mm wide on a vertical aerial photo. What is the flying height above the pavement for this photo if the camera focal length was 152.4 mm?

6-8. In the photo of Prob. 6-7, a rectangular building near the highway has photo dimensions of 6.1 mm and 9.2 mm. What is the actual size of the structure?

6-9. Repeat Prob. 6-8, except that in the photo of Prob. 6-7 a bridge appears. If its photo length is 28.9 mm, what is the actual length of the bridge?

6-10. A vertical photograph was taken, with a camera having a 152.9-mm focal length, from a flying height 2160 m above sea level. What is the scale of the photo at an elevation of 385 m above sea level? What is the datum scale?

6-11. Aerial photographs are to be taken for highway planning and design. If a 152-mm-focal-length camera is to be used and if an average scale of 1:6000 is required, what should be the flying height above average terrain?

6-12. A vertical aerial photograph was taken from a flying height of 3200 m above datum with a camera having a focal length of 209.45 mm. Highest, lowest, and average terrains appearing in the photograph are 1650 m, 1050 m, and 1370 m, respectively. Calculate the maximum, minimum, and average photographic scales.

6-13. A vertical photograph was taken over the lunar surface from an altitude of 72.5 km with a 80.20-mm-focal-length camera. What is the actual diameter of a crater whose diameter on the photograph scales 9.74 mm? (Give answer in km.)

6-14. Vertical photography for military reconnaissance is required. If the lowest safe flying altitude over enemy defenses is 5000 m, what camera focal length is necessary to achieve a photo scale of 1:24,000?

6-15. A distance *ab* on a vertical photograph is 65.0 mm, and the corresponding ground distance *AB* is 1153 m. If the camera focal length is 153.19 mm, what is the flying height above the terrain upon which line *AB* is located?

6-16. Vertical photography at an average scale of 1:5000 is to be acquired for the purpose of constructing a mosaic. What is the required flying height above average terrain if the camera focal length is 152.9 mm?

6-17. The distance on a map between two road intersections in flat terrain measures 43.6 mm. The distance between the same two points is 83.0 mm on a vertical photograph. If the scale of the map is 1:50,000, what is the scale of the photograph?

6-18. For Prob. 6-17, the intersections occur at an average elevation of 278 m above sea level. If the camera had a focal length of 209.6 mm, what is the flying height above sea level for this photo?

6-19. A section line scales 88.6 mm on a vertical aerial photograph. What is the scale of the photograph?

6-20. For Prob. 6-19, the average elevation of the section line is at 395 m above sea level, and the camera focal length is 152.4 mm. What would be the actual length of a ground line that lies at elevation 303 m above sea level and measures 53.1 mm on this photo?

6-21. A vertical aerial photo is exposed at 3910 ft above mean sea level using a camera having an 88.9-mm focal length. A triangular parcel of land that lies at elevation 850 ft above sea level appears on the photo, and its sides measure 39.5 mm, 28.9 mm, and 27.7 mm, respectively. What is the approximate area of this parcel in acres?

6-22. On a vertical aerial photograph, a line which was measured on the ground to be 536 m long scales 34.6 mm. What is the scale of the photo at the average elevation of this line?

6-23. Points *A* and *B* are at elevations 323 m and 422 m above datum, respectively. The photographic coordinates of their images on a vertical photograph are x_a = 68.27 mm, y_a = –32.37 mm, x_b = –87.44 mm, and y_b = 26.81 mm. What is the horizontal length of line *AB* if the photo was taken from 1535 m above datum with a 152.35-mm-focal-length camera?

6-24. Images *a*, *b*, and *c* of ground points *A*, *B*, and *C* appear on a vertical photograph taken from a flying height of 2625 m above datum. A 153.16-mm-focal-length camera was used. Points *A*, *B*, and *C* have elevations of 407 m, 311 m, and 379 m above datum, respectively. Measured photo coordinates of the images are x_a = –60.2 mm, y_a = 47.3 mm, x_b = 52.4 mm, y_b = 80.8 mm, x_c = 94.1 mm, and y_c = –79.7 mm. Calculate the horizontal lengths of lines *AB*, *BC*, and *AC* and the area within triangle *ABC* in hectares.

6-25. The image of a point whose elevation is 832 ft above datum appears 58.51 mm from the principal point of a vertical aerial photograph taken from a flying height of 6000 ft above datum. What would this distance from the principal point be if the point were at datum?

6-26. The images of the top and bottom of a utility pole are 113.6 mm and 108.7 mm, respectively, from the principal point of a vertical photograph. What is the height of the pole if the flying height above the base of the pole is 834 m?

6-27. An area has an average terrain elevation of 335 m above datum. The highest points in the area are 412 m above datum. If the camera focal plane opening is 23 cm square, what flying height above

datum is required to limit relief displacement with respect to average terrain elevation to 5.0 mm? (*Hint:* Assume the image of a point at highest elevation occurs in the corner of the camera format.) If the camera focal length is 152.7 mm, what is the resulting average scale of the photography?

6-28. The datum scale of a vertical photograph taken from 915 m above datum is $1{:}60\overline{00}$. The diameter of a cylindrical oil storage tank measures 6.69 mm at the base and 6.83 mm at the top. What is the height of the tank if its base lies at 213 m above datum?

6-29. Assume that the smallest discernible and measurable relief displacement that is possible on a vertical photo is 0.5 mm. Would it be possible to determine the height of a telephone utility box imaged in the corner of a 23-cm-square photo taken from 870 m above ground? (*Note:* Telephone utility boxes actually stand 1.2 m above the ground.)

6-30. If your answer to Prob. 6-29 is yes, what is the maximum flying height at which it would be possible to discern the relief displacement of the utility box? If your answer is no, at what flying height would the relief displacement of the box be discernible?

6-31. On a vertical photograph, images a and b of ground points A and B have photographic coordinates $x_a = -12.68$ mm, $y_a = 70.24$ mm, $x_b = 89.07$ mm, and $y_b = -92.41$ mm. The horizontal distance between A and B is 1317 m, and the elevations of A and B are 382 m and 431 m above datum, respectively. Calculate the flying height above datum if the camera had a 152.5-mm focal length.

6-32. Repeat Prob. 6-31, except that the horizontal distance AB is 1834 m and the camera focal length is 88.92 mm.

6-33. In Prob. 6-13, assume that the values given for focal length, photo distance, and flying height contain random errors of ±0.10 mm, ±0.05 mm, and ±0.30 km, respectively. What is the expected error in the computed diameter of the crater?

6-34. In Prob. 6-15, assume that the values given for focal length, photo distance, and ground length contain random errors of ±0.005 mm, ±0.50 mm, and ±0.30 m, respectively. What is the expected error in the computed flying height?

6-35. In Prob. 6-26, assume that the random error in each measured photo distance is ±0.10 mm and that the error in the flying height is ±2.0 m. What is the expected error in the computed height of the utility pole?

7-1 Depth Perception

In our daily activities we unconsciously measure depth or judge distances to a vast number of objects about us through our normal process of vision. Methods of judging depth may be classified as either *stereoscopic* or *monoscopic*. Persons with normal vision (those capable of viewing with both eyes simultaneously) are said to have *binocular* vision, and perception of depth through binocular vision is called stereoscopic viewing. *Monocular* vision is the term applied to viewing with only one eye, and methods of judging distances with one eye are termed *monoscopic*. A person having normal binocular vision can, of course, view monocularly by covering one eye.

Distances to objects, or depths, can be perceived monoscopically on the basis of (1) relative sizes of objects, (2) hidden objects, (3) shadows, and (4) differences in focusing of the eye required for viewing objects at varying distances. Examples of the first two of these are shown in Fig. 7-1. Depth to the far end of the football field may be perceived, for example, on the basis of the relative sizes of the goalposts. The goalposts are actually the same size, of course, but one appears smaller because it is farther away. Also, the stands at the far end of the stadium are quickly judged to be a considerable distance away from the goalposts in the foreground because they are partially obstructed by the goalposts.

FIGURE 7-1 Depth perception by relative sizes and hidden objects.

Monoscopic methods of depth perception enable only rough impressions to be gained of distances to objects. With stereoscopic viewing, however, a much greater degree of accuracy in depth perception can be attained. Stereoscopic depth perception is of fundamental importance in photogrammetry, for it enables the formation of a three-dimensional stereomodel by viewing a pair of

overlapping photographs. The stereomodel can then be studied, measured, and mapped. An explanation of how this phenomenon is achieved is the subject of this chapter, and explanations of its use in measuring and mapping are given in the chapters that follow.

7-2 The Human Eye

The phenomenon of stereoscopic depth perception can be more clearly understood with the help of a brief description of the anatomy and physiology of the human eye. The human eye functions in much the same manner as a camera. As shown in Fig. 7-2, the eye is essentially a spherical organ having a circular opening called the *pupil*. The pupil is protected by a transparent coating called the *cornea*. Incident light rays pass through the cornea, enter the eye through the pupil, and strike the *lens*, which is directly behind the pupil. The cornea and lens refract the light rays according to Snell's law (see Sec. 2-2).

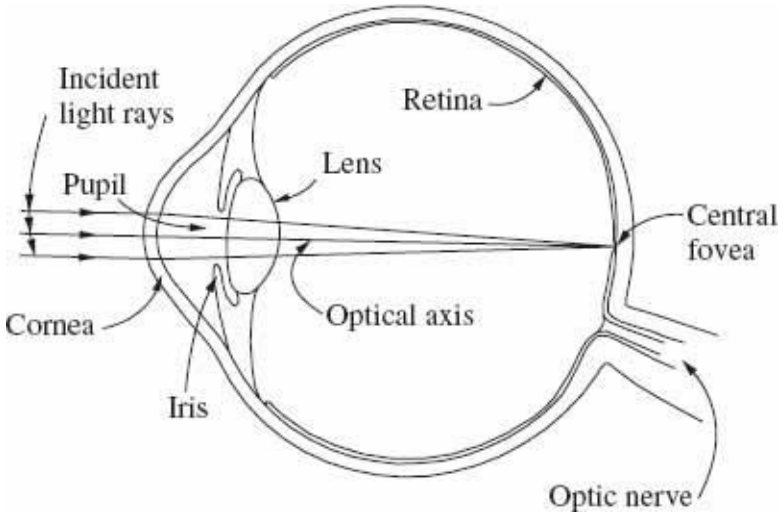


FIGURE 7-2 Cross section of the human eye.

The lens of the eye is biconvex and is composed of a refractive transparent medium. It is suspended by many muscles, which enable the lens to be moved so that the *optical axis* of the eye can be aimed directly at an object to be viewed. As with a camera, the eye must satisfy the lens formula, Eq. (2-4), for each different object distance. The eye's image distance is constant, however; therefore, to satisfy the lens formula for varying object distances, the focal length of the lens changes. When a distant object is viewed, the lens muscles relax, causing the spherical surfaces of the lens to become flatter. This increases the focal length to satisfy the lens formula and accommodate the long object distance. When close objects are viewed, a reverse procedure occurs. The eye's ability to focus for varying object distances is called *accommodation*.

As with a camera, the eye has a diaphragm called the *iris*. The iris (colored part of the eye) automatically contracts or expands to regulate the amount of light entering the eye. When the eye is subjected to intense light, the iris contracts, reducing the pupil aperture. When the intensity of light lessens, the iris dilates to admit more light.

The cornea partially refracts incident light rays before they encounter the lens. The lens refracts them further and brings them to focus on the *retina*, thereby forming an image of the viewed object. The retina is composed of very delicate tissue. The most important region of the retina is the *central fovea*, a small pit near the intersection of the optical axis with the retina. The central fovea is the area

of sharpest vision. The retina performs a function similar to that performed by the emulsion of photographic film. When it is stimulated by light, the sense of vision is caused, which is transmitted to the brain via the *optic nerve*.

7-3 Stereoscopic Depth Perception

With binocular vision, when the eyes fixate on a certain point, the optical axes of the two eyes converge on that point, intersecting at an angle called the *parallactic angle*. The nearer the object, the greater the parallactic angle, and vice versa. In Fig. 7-3, the optical axes of the two eyes L and R are separated by a distance b_e, called the *eye base*. For the average adult, this distance is between 63 and 69 mm, or approximately 2.6 in. When the eyes fixate on point A, the optical axes converge, forming parallactic angle ϕ_a. Similarly, when sighting an object at B, the optical axes converge, forming parallactic angle ϕ_b. The brain automatically and unconsciously associates distances D_A and D_B with corresponding parallactic angles ϕ_a and ϕ_b. The depth between objects A and B is $D_B - D_A$ and is perceived from the difference in these parallactic angles.

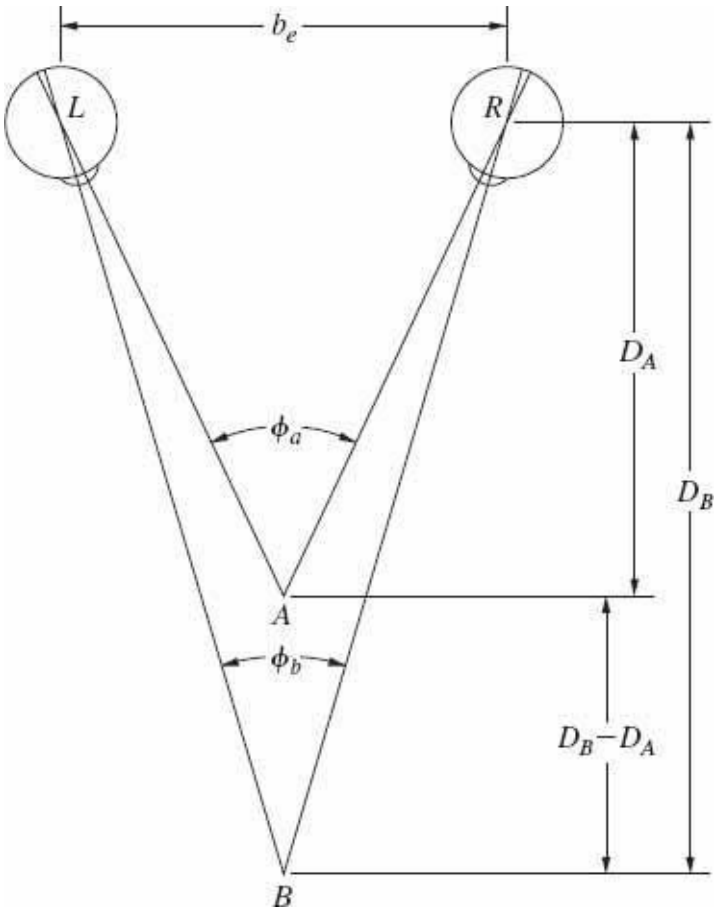


FIGURE 7-3 Stereoscopic depth perception as a function of parallactic angle.

The ability of human beings to detect changes in parallactic angles, and thus judge differences in depth, is quite remarkable. Although it varies somewhat among individuals, the average person is capable of discerning parallactic angle changes of about 3 seconds of arc, but some are able to perceive changes as small as 1 second of arc. This means that photogrammetric procedures for determining heights of objects and terrain variations based on depth perception by comparisons of parallactic angles can be highly precise.

7-4 Viewing Photographs Stereoscopically

Suppose that while a person is gazing at object *A* of Fig. 7-4, a transparent medium containing image marks a_1 and a_2 is placed in front of the eyes as shown. Assume further that the image marks are identical in shape to object *A*, and that they are placed on the optical axes so that the eyes are unable to detect whether they are viewing the object or the two marks. Object *A* could therefore be removed without any noticeable changes in the images received on the retinas of the eyes. As shown in Fig. 7-4, if the image marks are moved closer together to, say, a_1', and a_2', the parallactic angle increases and the object is perceived to be nearer the eyes at *A*′. If the marks are moved farther apart to a_1'' and a_2'', the parallactic angle decreases and the brain receives an impression that the object is farther away, at *A*″.

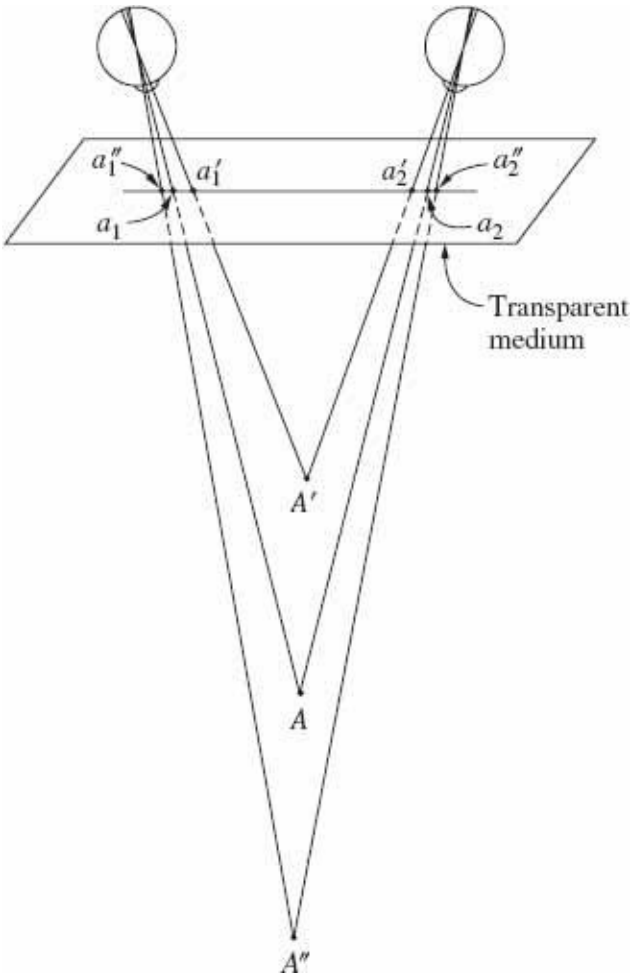


FIGURE 7-4 The apparent depth to the object A can be changed by changing the spacing of the images.

The phenomenon of creating the three-dimensional or stereoscopic impression of objects by viewing identical images of the objects can be achieved photographically. Suppose that a pair of aerial photographs is taken from exposure stations L_1 and L_2 so that the building appears on both photos, as shown in Fig. 7-5. Flying height above ground is *H*′, and the distance between the two exposures is *B*, the *air base*. Object points *A* and *B* at the top and bottom of the building are imaged at a_1 and b_1 on the left photo and at a_2 and b_2 on the right photo. Now, if the two photos are laid on a table and viewed so that the left eye sees only the left photo and the right eye sees only the right photo, as shown in Fig. 7-6, a three-dimensional impression of the building is obtained. The three-dimensional impression

appears to lie below the tabletop at a distance h from the eyes. The brain judges the height of the building by associating depths to points A and B with the parallactic angles ϕ_a and ϕ_b, respectively. When the eyes gaze over the entire overlap area, the brain receives a continuous three-dimensional impression of the terrain. This is achieved by the continuous perception of changing parallactic angles of the infinite number of image points which make up the terrain. The three-dimensional model thus formed is called a *stereoscopic model* or simply a *stereomodel*, and the overlapping pair of photographs is called a *stereopair*.

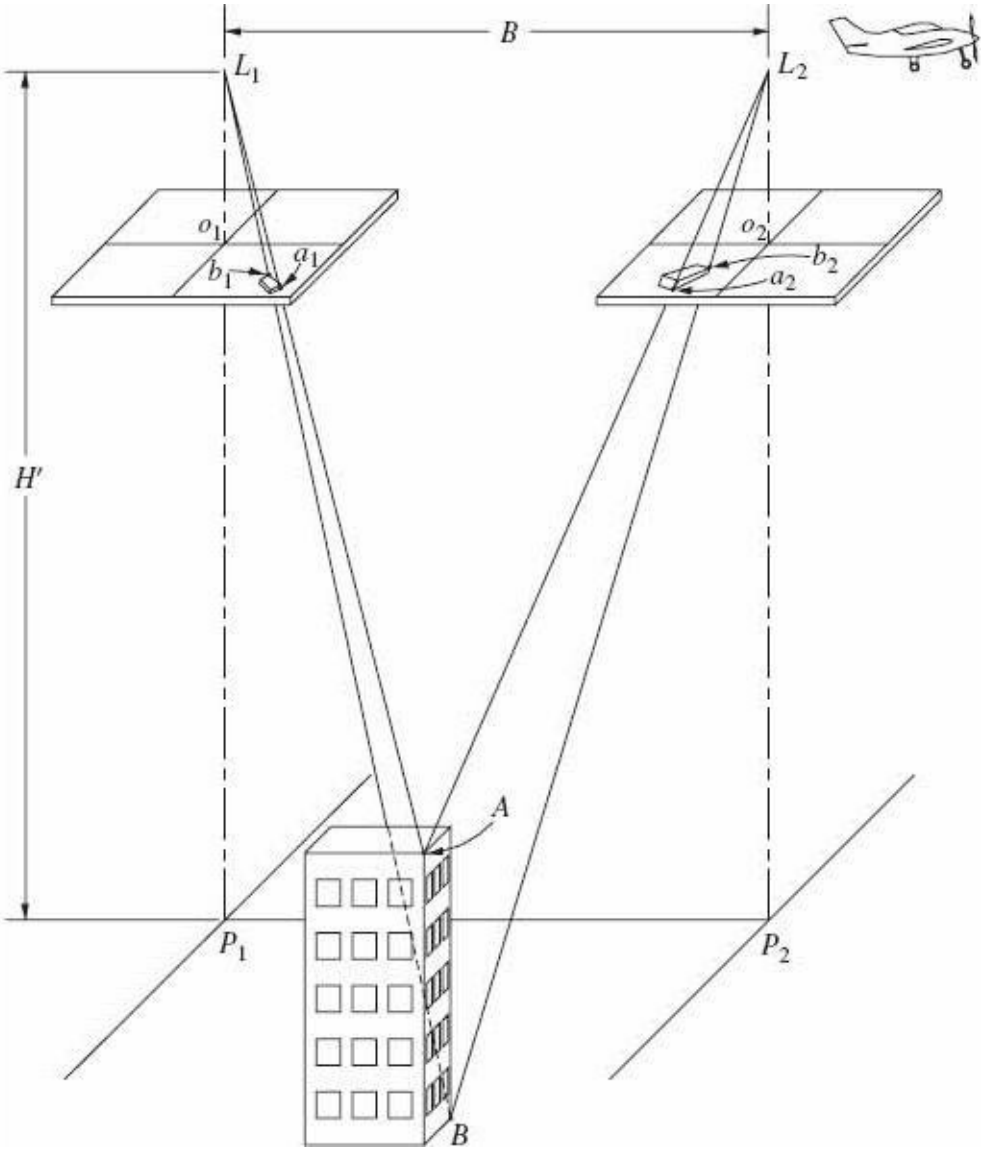


FIGURE 7-5 Photographs from two exposure stations with building in common overlap area.

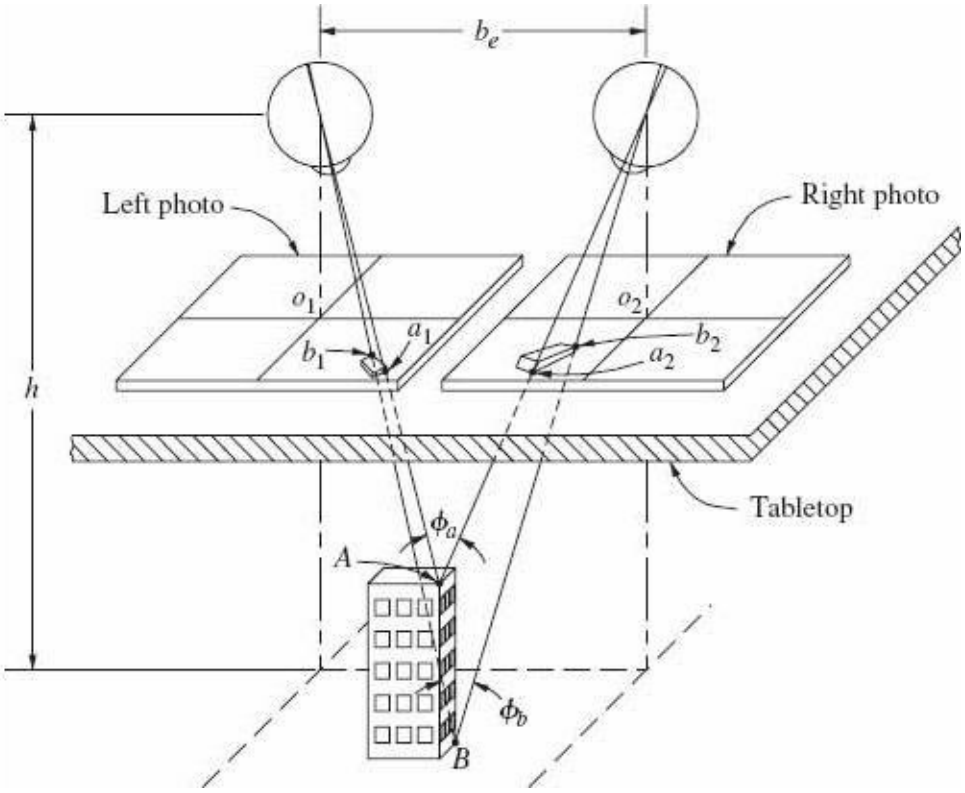


FIGURE 7-6 Viewing the building stereoscopically.

7-5 Stereoscopes

It is quite difficult to view photographs stereoscopically without the aid of optical devices, although some individuals can do it. Besides being an unnatural operation, one of the major problems associated with stereoviewing without optical aids is that the eyes are focused on the photos, while at the same time the brain perceives parallactic angles which tend to form the stereomodel at some depth beyond the photos—a confusing situation, to say the least. These difficulties in stereoscopic viewing may be overcome through the use of instruments called *stereoscopes*.

There is a wide selection of stereoscopes serving a variety of special purposes. All operate in essentially the same manner. The *lens* or *pocket* stereoscope, shown in Fig. 7-7, is the least expensive and most commonly used stereoscope. It consists of two simple convex lenses mounted on a frame. The spacing between the lenses can be varied to accommodate various eye bases. The legs fold or can be removed so that the instrument is easily stored or carried—a feature which renders the pocket stereoscope ideal for fieldwork. A schematic diagram of the pocket stereoscope is given in Fig. 7-8. The legs of the pocket stereoscope are slightly shorter than the focal length f of the lenses. When the stereoscope is placed over the photos, light rays emanating from points such as a_1 and a_2 on the photos are refracted slightly as they pass through each lens. (Recall from Chap. 2 that a bundle of light rays from a point exactly at a distance f from a lens will be refracted and emerge through the lens parallel.) The eyes receive the refracted rays (shown dashed in Fig. 7-8), and on the basis of the eye focusing associated with these incoming rays, the brain receives the impression that the rays actually originate from a greater distance than that to the tabletop upon which the photos rest. This overcomes the difficulties noted above. The lenses also serve to magnify the images, thereby enabling details to be seen more clearly.

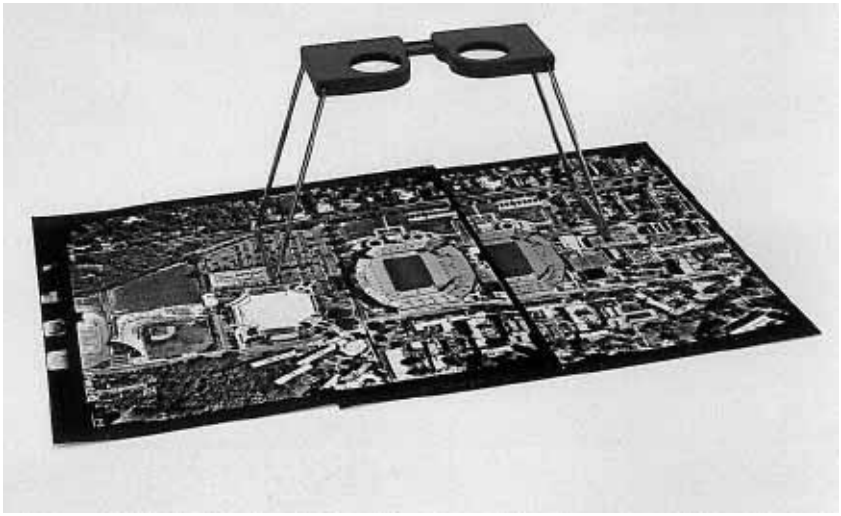


FIGURE 7-7 Lens or pocket stereoscope. (*Courtesy University of Florida.*)

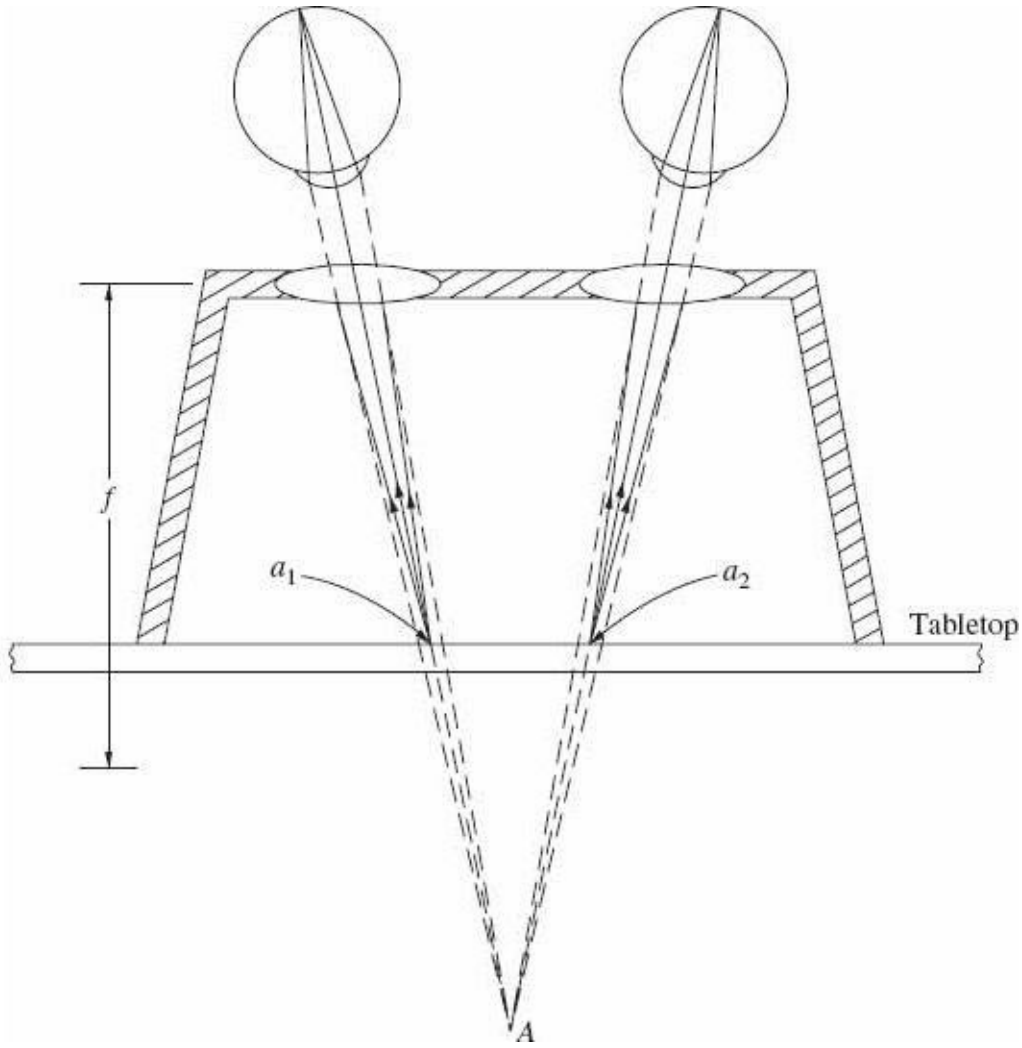


FIGURE 7-8 Schematic diagram of the pocket stereoscope.

In using a pocket stereoscope, the photos are placed so that corresponding images are slightly less than the eye base apart, usually about 5 cm. For normal 23-cm-format photos taken with 60 percent end lap, the common overlap area of a pair of photos is a rectangular area about 14 cm wide, as shown crosshatched in Fig. 7-9*a*. If the photos are separated by 5 cm for stereoviewing with a pocket stereoscope, as shown in Fig. 7-9*b*, there is a rectangular area, shown double crosshatched, in which the top photo obscures the bottom photo, thereby preventing stereoviewing. To overcome this

problem, the top photo can be gently rolled up out of the way to enable viewing the corresponding imagery of the obscured area.

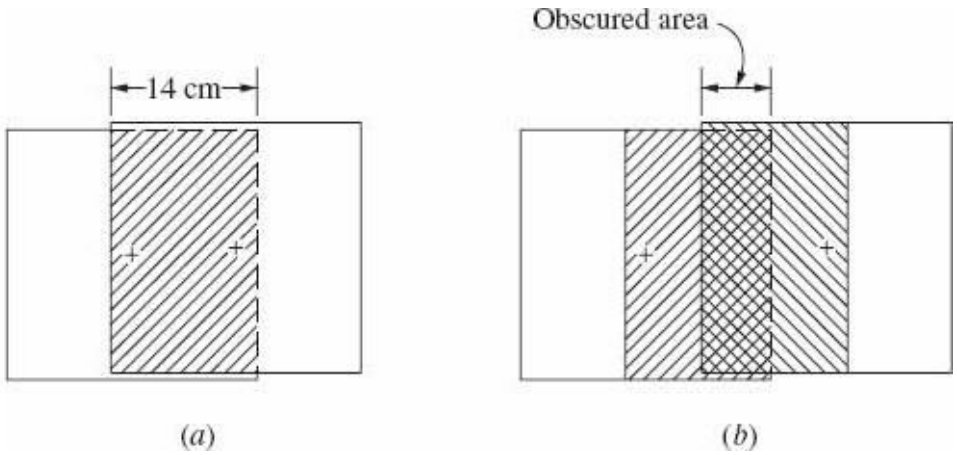


FIGURE 7-9 (*a*) The common overlap area of a pair of 23-cm-format photos taken with 60 percent end lap (corresponding images coincident). (*b*) Obscured area when photos are oriented for viewing with pocket stereoscope.

The *mirror* stereoscope shown in Fig. 7-10 permits the two photos to be completely separated when viewed stereoscopically. This eliminates the problem of one photo obscuring part of the overlap of the other, and it also enables the entire width of the stereomodel to be viewed simultaneously. The operating principle of the mirror stereoscope is illustrated in Fig. 7-11. The stereoscope has two large wing mirrors and two smaller eyepiece mirrors, all of which are mounted at 45° to the horizontal. Light rays emanating from image points on the photos such as a_1 and a_2 are reflected from the mirror surfaces, according to the principles of reflection discussed in Sec. 2-2, and are received at the eyes, forming parallactic angle ϕ_a. The brain automatically associates the depth to point A with that parallactic angle. The stereomodel is thereby created beneath the eyepiece mirrors, as illustrated in Fig. 7-11.

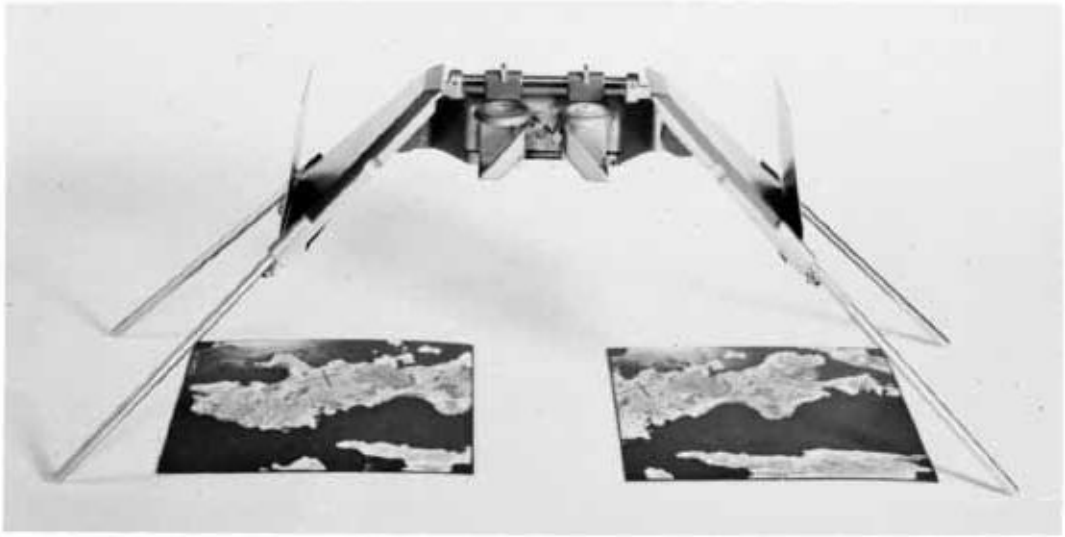


FIGURE 7-10 ST-4 mirror stereoscope. (*Courtesy LH Systems, LLC.*)

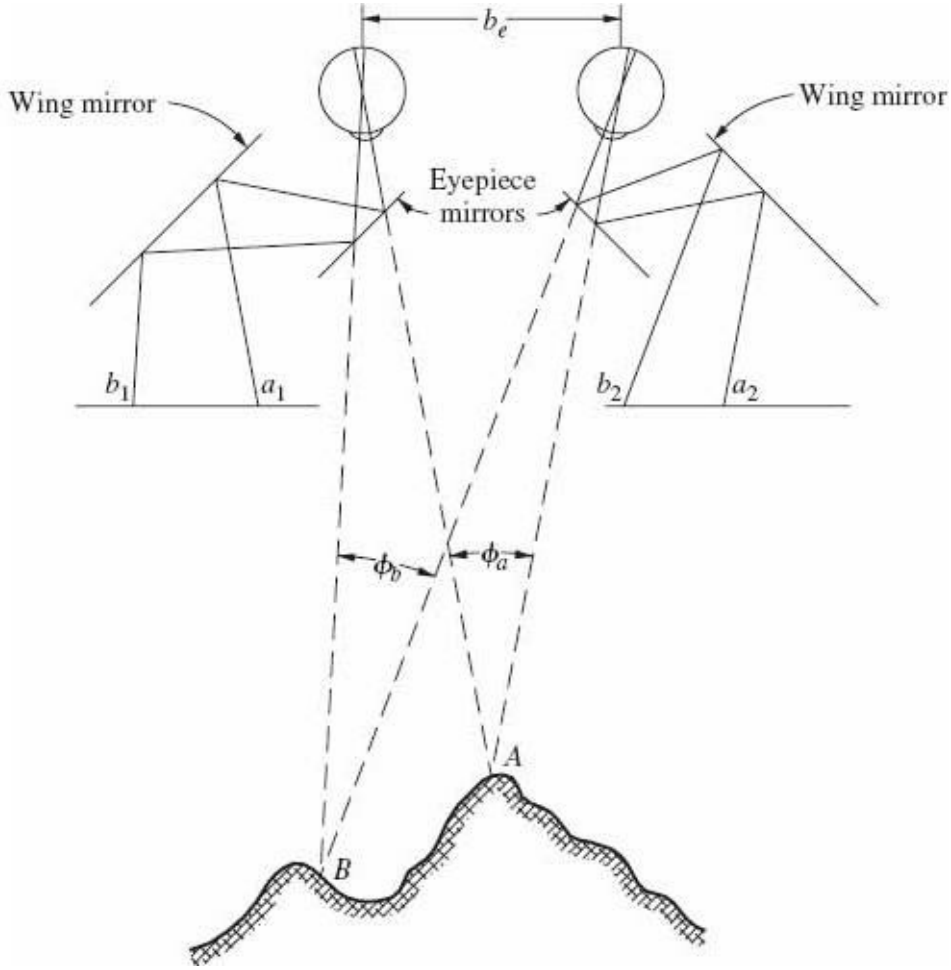


FIGURE 7-11 Operating principle of the mirror stereoscope.

Simple lenses are usually placed directly above the eyepiece mirrors as shown in Fig. 7-10. Their separation may be changed to accommodate various eye bases. The focal length of these lenses is also slightly greater than the length of the reflected ray path from photo to eyes, and therefore they serve basically the same function as the lenses of the pocket stereoscope. Mirror stereoscopes may be equipped with binoculars which fasten over the eyepiece mirrors. The binoculars, which may be focused individually to accommodate each eye, permit viewing images at high magnification—a factor which is especially important and useful in photo interpretation or in identifying small image points. High magnification, of course, limits the field of view so that the entire stereomodel cannot be viewed simultaneously. The stereoscope must therefore be moved about if all parts of the stereomodel are to be seen.

It is extremely important to avoid touching the first-surfaced mirrors of a mirror stereoscope. This is true because the hands contain oils and acids which can tarnish the coatings on the mirrors, rendering them useless. If the mirrors are accidently soiled by fingerprints, they should be cleaned immediately, using a soft cloth and lens-cleaning fluid.

A different type of stereoscope called the *zoom stereoscope* is shown in Fig. 7-12. A variety of these instruments are manufactured, affording a choice of special features such as continuous zoom magnification up to 120X, capability of rotating images optically (which permits convenient correction for crab or alignment), and differential enlargement so that two photos of different scales can be viewed stereoscopically. For direct stereoscopic viewing of film negatives, these stereoscopes may be obtained mounted on a light table and equipped with a special scanning mechanism. A reel of film and a take-up reel are mounted on either end of the table. By turning a crank, the frames are brought into position for viewing.

FIGURE 7-12 Zoom 95 stereoscope. (*Courtesy Bausch and Lomb Co.*)

Other systems have been developed to facilitate stereoscopic viewing. These systems employ schemes such as colored filters, polarized light, or alternating imagery in order to achieve the stereoscopic effect. Sections 12-5 and 12-15 include a description of several of these approaches as they apply to stereoscopic plotting instruments.

7-6 The Use of Stereoscopes

Before you attempt to use a stereoscope, it is important to study the operator's manual if one is available. This is especially true for stereoscopes having more elaborate optical viewing systems. Also, the lenses and mirrors should be inspected and cleaned if necessary.

In stereoscopic viewing, it is important to orient the photos so that the left and right eyes see the left and right photos, respectively. If the photos are viewed in reverse, a *pseudoscopic view* results in which ups and downs are reversed; e.g., valleys appear as ridges and hills appear as depressions. This can be advantageous for certain work such as tracing drainage patterns, but normally the correct stereoscopic view is desired.

Accurate and comfortable stereoscopic viewing requires that the eye base, the line joining the centers of the stereoscope lenses, and the flight line all be parallel. Therefore, after the photos have been inspected and laid out so as to prevent a pseudoscopic view, the flight line is marked on both photos. For vertical photographs, the flight line is the line from the center of the left photo to the center of the right photo. In marking the flight line, the photo centers (principal points) are first located by joining opposite fiducial marks with straight lines. Principal points are shown at o_1 and o_2 on Fig. 7-13. *Corresponding principal points* (also called *conjugate principal points*), which are the locations of principal points on adjacent overlapping photos, are marked next. This may be done satisfactorily by carefully observing images immediately surrounding the principal points, and then marking the corresponding principal points by estimating their positions with respect to these surrounding images. If the area is level and there are identifiable features in the photos, the method

illustrated in Fig. 7-14 can be used to find the conjuagte principal point with acceptable accuracy. In Fig. 7-14*a*, intersecting lines joining opposite fiducial marks define the principal point of the left photo. Since there is no distinct feature located at this intersection, distances R_1 and R_2 are measured to nearby features at approximately the same elevation. Figure 7-14*b* shows the corresponding area on the right photo, where arcs centered on these features and having radii R_1 and R_2 are intersected to obtain the conjugate principal point. The corresponding principal points are shown at o_1' and o_2' on Fig. 7-13.

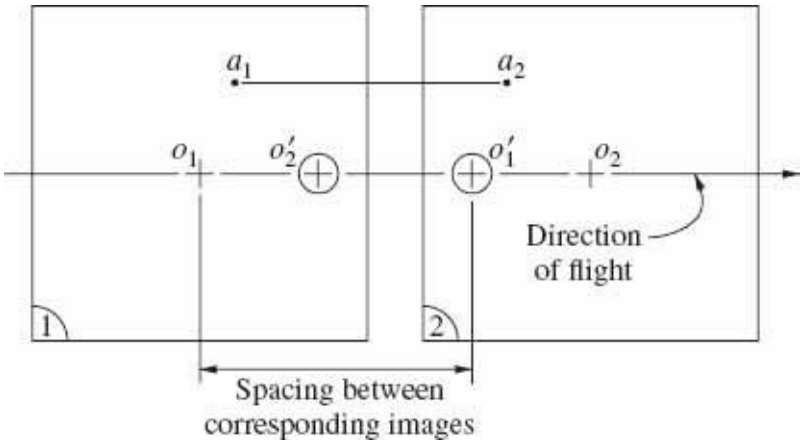
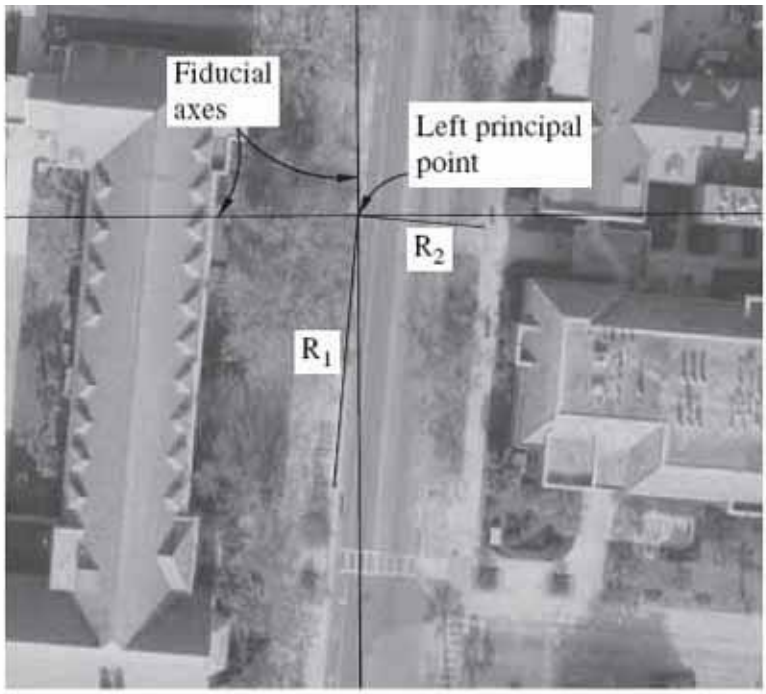
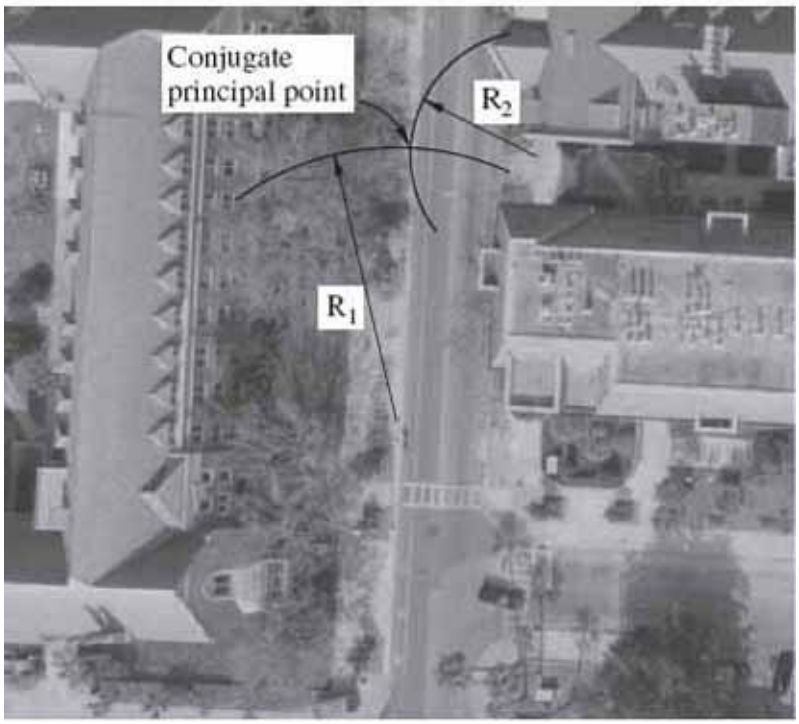


FIGURE 7-13 Pair of photos properly oriented for stereoscopic viewing.



(*a*)



(*b*)

FIGURE 7-14 (*a*) Center portion of left photo showing its principal point and distances to two manhole covers. (*b*) Intersection of corresponding distances to the same features at the conjugate principal point in the right photo.

The next step in orienting a pair of photos for stereoscopic viewing is to fasten the left photo down onto the table. Then the right photo is oriented so that the four points defining the flight line (o_1, o_2', o_1', and o_2) all lie along a straight line, as shown in Fig. 7-13. The right photo is retained in this orientation, and while being viewed through the stereoscope, it is moved sideways until the spacing between corresponding images produces a comfortable stereoscopic view. Normally the required spacing between corresponding images is slightly more than 5 cm for a pocket stereoscope and about 25 cm for a mirror stereoscope.

It is not absolutely necessary to mark the flight lines and orient photos for stereoscopic viewing in the manner outlined above; in fact, for casual stereoviewing, the geometry shown in Fig. 7-13 is normally achieved by a trial method in which the photos are simply shifted in position until a clear stereoscopic view is obtained. If accuracy and eye comfort are considerations, however, orientation by the flight-line procedure is recommended.

As previously stated, comfortable stereoscopic viewing requires that the line joining the stereoscope lens centers be parallel with the flight line. Once the photos are properly oriented, the operator can easily align the stereoscope by simply rotating it slightly until the most comfortable viewing position is obtained. The operator should look directly into the centers of the lenses, thereby holding the eye base parallel with the flight line.

7-7 Causes of *Y* Parallax

An essential condition which must exist for clear and comfortable stereoscopic viewing is that the line joining corresponding images be parallel with the direction of flight. This condition is fulfilled with the corresponding images a_1 and a_2 shown in Fig. 7-13. When corresponding images fail to lie along a line parallel to the flight line, *y parallax*, denoted by p_y, is said to exist. Any slight amount of *y* parallax causes eyestrain, and excessive amounts prevent stereoscopic viewing altogether.

If a pair of truly vertical overlapping photos taken from equal flying heights is oriented perfectly, then no *y* parallax should exist anywhere in the overlap area. Failure of any of these conditions to be satisfied will cause *y* parallax. In Fig. 7-15, for example, the photos are improperly oriented, and the principal points and corresponding principal points do not lie on a straight line. As a result, *y* parallax exists at both points *a* and *b*. This condition can be prevented by careful orientation.

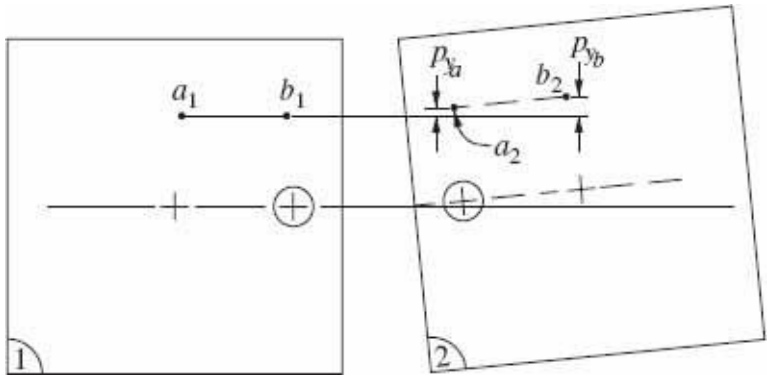


FIGURE 7-15 Here *y* parallax is caused by improper orientation of the photos.

In Fig. 7-16 the left photo was exposed from a lower flying height than the right photo, and consequently its scale is larger than the scale of the right photo. Even though the photos are truly vertical and properly oriented, *y* parallax exists at both points *a* and *b* due to variation in flying heights. To obtain a comfortable stereoscopic view, the *y* parallax can be eliminated by sliding the right photo upward transverse to the flight line when viewing point *a* and sliding it downward when viewing point *b*.

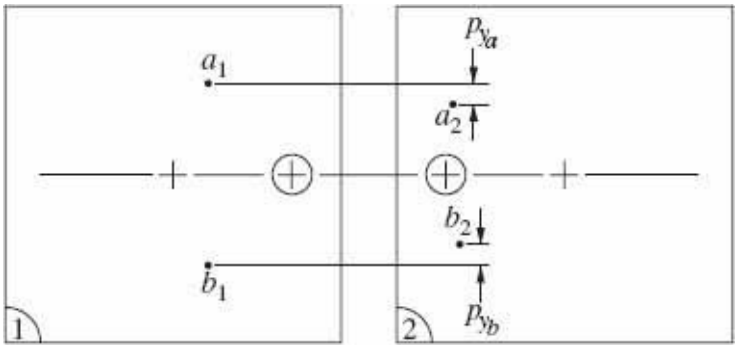


FIGURE 7-16 The *y* parallax is caused by variation in flying height.

The effect of tilted photos is illustrated in Fig. 7-17. The left photo is truly vertical and shows positions of images *a* through *d* of a square parcel of property on flat terrain. The right photo was tilted such that the same parcel appears as a trapezoid. In this case, *y* parallax exists throughout the stereoscopic model as a result of the tilt, as indicated for points *a* and *c*. In practice, the direction of tilt is random, and therefore small *y* parallaxes from this source are likely to exist in variable amounts throughout most stereomodels. If it is the intent to obtain vertical photography from a constant flying height, however, these conditions are generally so well controlled that *y* parallaxes from these sources are seldom noticeable. Most serious *y* parallaxes usually occur from improper orientation of the photos, a condition which can be easily corrected.

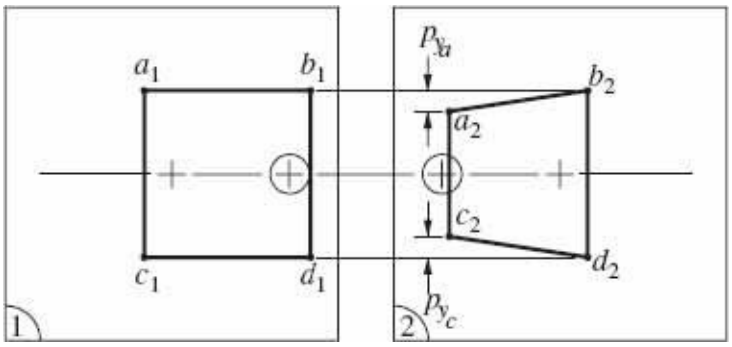


FIGURE 7-17 The *y* parallax is caused by tilt of the photos.

7-8 Vertical Exaggeration in Stereoviewing

Under normal conditions, the vertical scale of a stereomodel will appear to be greater than the horizontal scale; i.e., an object in the stereomodel will appear to be too tall. This apparent scale disparity is called *vertical exaggeration*. It is usually of greatest concern to photo interpreters, who must take this condition into account when estimating heights of objects, rates of slopes, etc.

Although other factors are involved, vertical exaggeration is caused primarily by the lack of equivalence of the *photographic base-height ratio, B/H′*, and the corresponding *stereoviewing base-height ratio, b_e/h*. The term *B/H′* is the ratio of the *air base* (distance between the two exposure stations) to flying height above average ground, and b_e/h is the ratio of the *eye base* (distance between the two eyes) to the distance from the eyes at which the stereomodel is perceived. Figures 7-18*a* and *b* depict, respectively, the taking of a pair of vertical overlapping photographs and the stereoscopic viewing of those photos. In Fig. 7-18*a*, the camera focal length is *f*, the air base is *B*, the flying height above ground is *H′*, the height of ground object *AC* is *Z*, and the horizontal ground distance *KC* is *D*. In Fig. 7-18*a*, assume that *Z* is equal to *D*. In Fig. 7-18*b*, *i* is the image distance from the eyes to the

photos, b_e is the eye base, h is the distance from the eyes to the perceived stereomodel, z is the stereomodel height of object $A'C'$, and d is the horizontal stereomodel distance $K'C'$. Note that while the ratio Z/D is equal to 1, the ratio z/d is greater than 1 due to vertical exaggeration.

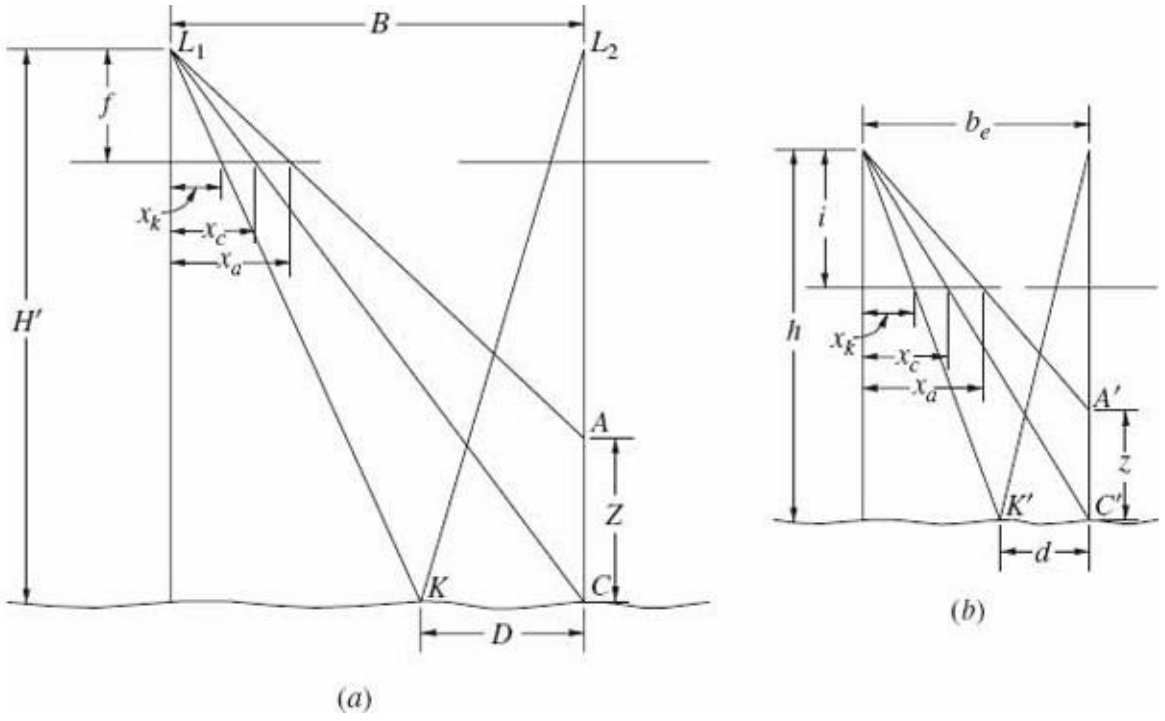


FIGURE 7-18 Simplistic diagrams for analyzing vertical exaggeration. (*a*) Geometry of overlapping aerial photography. (*b*) Geometry of stereoscopic viewing of the photos of part (a).

An equation for calculating vertical exaggeration can be developed with reference to these figures. From similar triangles of Fig. 7-18*a*,

$$\frac{x_a}{B} = \frac{f}{H' - Z} \qquad \text{from which} \qquad x_a = \frac{Bf}{H' - Z} \tag{a}$$

$$\text{Also} \qquad \frac{x_c}{B} = \frac{f}{H'} \qquad \text{from which} \qquad x_c = \frac{Bf}{H'} \tag{b}$$

Subtracting (*b*) from (*a*) and reducing gives

$$x_a - x_c = Bf \frac{Z}{(H')^2 - H'Z} \tag{c}$$

Also from similar triangles of Fig. 7-18*b*,

$$\frac{x_a}{b_e} = \frac{i}{h - z} \qquad \text{from which} \qquad x_a = \frac{b_e i}{h - z} \tag{d}$$

$$\text{and} \qquad \frac{x_c}{b_e} = \frac{i}{h} \qquad \text{from which} \qquad x_c = \frac{b_e i}{h} \tag{e}$$

Subtracting (*e*) from (*d*) and reducing yields

$$x_a - x_c = b_e i \frac{z}{h^2 - hz} \tag{f}$$

Equating (*c*) and (*f*) gives

$$Bf \frac{Z}{(H')^2 - H'Z} = b_e i \frac{z}{h^2 - hz}$$

In the above equation, the values of Z and z are normally considerably smaller than the values of H' and h, respectively; thus

$$\frac{BfZ}{(H')^2} \approx \frac{b_e i z}{h^2} \qquad \text{from which} \qquad \frac{z}{Z} = \frac{fh}{H'i} \frac{Bh}{H'b_e} \tag{g}$$

Also from similar triangles of Figs. 7-18*a* and *b*,

$$\frac{x_c - x_k}{D} = \frac{f}{H'} \qquad \text{from which} \qquad D = (x_c - x_k) = \frac{H'}{f} \tag{h}$$

$$\text{and} \qquad \frac{x_c - x_k}{d} = \frac{i}{h} \qquad \text{from which} \qquad d = (x_c - x_k)\frac{h}{i} \tag{i}$$

Dividing (*i*) by (*h*) and reducing yields

$$\frac{d}{D} = \frac{fh}{H'i} \tag{j}$$

Substituting (*j*) into (*g*) and reducing gives

$$\frac{z}{Z} = \frac{d}{D} \frac{Bh}{H'b_e} \tag{k}$$

In Eq. (*k*), if the term $Bh/(H'b_e)$ is equal to 1, there is no vertical exaggeration of the stereomodel. (Recall that Z is equal to D.) Thus an expression for the magnitude of vertical exaggeration V is given by

$$V \approx \frac{B}{H'} \frac{h}{b_e} \tag{7-1}$$

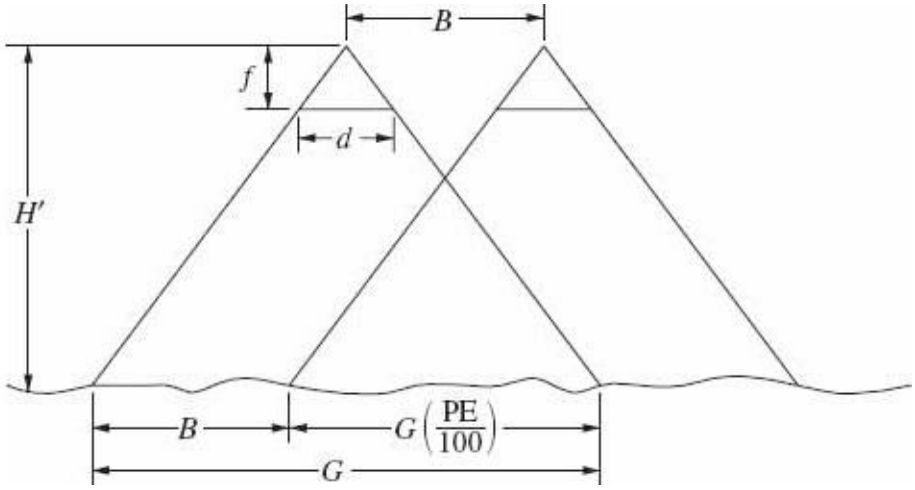


FIGURE 7-19 Base-height ratio (B/H').

From Eq. (7-1) it is seen that the magnitude of vertical exaggeration in stereoscopic viewing can be approximated by multiplying the B/H' ratio by the inverse of the b_e/h ratio. An expression for the B/H' ratio can be developed with reference to Fig. 7-19. In this figure, G represents the total ground coverage of a vertical photo taken from an altitude of H' above ground. Air base B is the distance between exposures. From the figure,

$$B = G - G\frac{PE}{100} = G\left(1 - \frac{PE}{100}\right) \tag{l}$$

In Eq. (*l*), PE is the percentage of end lap, which gives the amount that the second photo overlaps the first. Also by similar triangles of the figure,

$$\frac{H'}{G} = \frac{f}{d} \qquad \text{from which} \qquad H' = \frac{fG}{d} \tag{m}$$

In Eq. (*m*), f is the camera focal length and d its format dimension. Dividing Eq. (*l*) by Eq. (*m*) and reducing gives

$$\frac{B}{H'} = \left(1 - \frac{PE}{100}\right)\frac{d}{f} \tag{7-2}$$

The stereoviewing base-height ratio varies due to differences in the distances between the eyes of users and varying dimensions of stereoscopes. It can, however, be approximated in the following way. Figure 7-20 illustrates the relationships involved in this approximation. With an eye base, b_e, averaging about 65 mm in humans, we need only to find the perceived distance from the eyes to the stereomodel, h to make an approximation. If the distance between the photos is b_s, and the distance of the stereoscope from the photos is i, then we can use the following equation to estimate h by similar triangles:

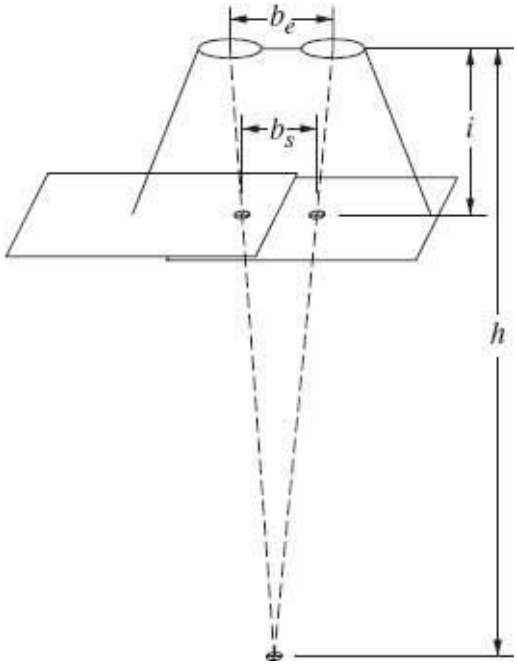


FIGURE 7-20 Eye base to perceived model height ratio.

$$h = i\frac{b_e}{b_e - b_s} \tag{n}$$

Equation (*n*) can be used to form Eq. (7-3) to directly solve for the stereoviewing base-height ratio:

$$\frac{b_e}{h} = \frac{b_e - b_s}{i} \tag{7-3}$$

Example 7-1
Estimate the stereoviewing base-height ratio if the height of the stereoscope above the photos is 10.00 cm, and the photos are placed 5.0 cm apart.

Solution By Eq. (7-3)

$$\frac{b_e}{h} = \frac{b_e - b_s}{i} = \frac{(6.5\text{ cm} - 5.0\text{ cm})}{10.0\text{ cm}} = 0.15 \quad \blacktriangle$$

Example 7-2
Using the results of Example 7-1, calculate the approximate vertical exaggeration for vertical aerial photos taken with a 152.4-mm-focal-length camera having a 23-cm-square format if the photos were taken with 60 percent end lap.

Solution By Eq. (7-2),

$$\frac{B}{H'} = \left(1 - \frac{60}{100}\right)\frac{230}{152.4} = 0.60$$

From Example 7-1, b_e/h is approximately 0.15,

$$V = 0.60\left(\frac{1}{0.15}\right) = 4.0 \qquad \text{(approx.)}$$

Note: If a 305-mm-focal-length camera had been used, the B/H' ratio would have been 0.30, and vertical exaggeration would have been reduced to 2. ▲

References

Ambrose, W. R.: "Stereoscopes with High Performance," *Photogrammetric Engineering*, vol. 31, no. 5, 1965, p. 822.

American Society of Photogrammetry: *Manual of Photogrammetry*, 4th ed., Bethesda, MD, 1980, chap. 10.

Collins, S. H.: "Stereoscopic Depth Perception," *Photogrammetric Engineering and Remote Sensing*, vol. 47, no. 1, 1981, p. 45.

Dalsgaard, J.: "Stereoscopic Vision—A Problem in Terrestrial Photogrammetry," *Photogrammetria*, vol. 34, no. 1, 1978, p. 3.

El Hassan, I. A.: "A New Formula for Vertical Exaggeration in Stereo Models," Map Asia and ISG 2010, 26-28 July 2010, Kuala Lumpur, Malaysia.

Gumbel, E. J.: "The Effect of the Pocket Stereoscope on Refractive Anomalies of the Eyes," *Photogrammetric Engineering*, vol. 30, no. 5, 1964, p. 795.

Howard, A. D.: "The Fichter Equation for Correcting Stereoscopic Slopes," *Photogrammetric Engineering*, vol. 34, no. 4, 1968, p. 386.

LaPrade, G. L.: "Stereoscopy—A More General Theory," *Photogrammetric Engineering*, vol. 38, no. 12, 1972, p. 1177.

———: "Stereoscopy—Will Dogma or Data Prevail?" *Photogrammetric Engineering*, vol. 39, no. 12, 1973, p. 1271.

Miller, C. I.: "Vertical Exaggeration in the Stereo Space Image and Its Use," *Photogrammetric Engineering*, vol. 26, no. 5, 1960, p. 815.

Myers, B. J., and F. P. Van der Duys: "A Stereoscopic Field Viewer," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 12, 1975, p. 1477.

Nicholas, G., and J. T. McCrickerd: "Holography and Stereoscopy: The Holographic Stereogram," *Photographic Science and Engineering*, vol. 13, no. 6, 1969, p. 342.

Palmer, D. A.: "Stereoscopy and Photogrammetry," *Photogrammetric Record*, vol. 4, 1964, p. 391.

Raju, A. V., and E. Parthasarathi: "Stereoscopic Viewing of Landsat Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 10, 1977, p. 1243.

Scheaffer, C. E.: "Stereoscope for Strips," *Photogrammetric Engineering*, vol. 34, no. 10, 1968, p. 1044.

Thayer, T. P.: "The Magnifying Single Prism Stereoscope: A New Field Instrument," *Journal of Forestry*, vol. 61, 1963, p. 381.

Yacoumelos, N.: "The Geometry of the Stereomodel," *Photogrammetric Engineering*, vol. 38, no. 8, 1972, p. 791.

Problems

7-1. What are some of the monocular methods of perceiving depth?

7-2. What is a parallactic angle?

7-3. Compare the advantages and disadvantages of the pocket and mirror stereoscopes.

7-4. Give a step-by-step procedure for orienting photos for stereoscopic viewing.

7-5. What is *y* parallax? What are the causes of *y* parallax in a stereomodel?

7-6. Prepare a table of B/H' ratios for camera focal lengths of 91, 152, 213, and 316 mm; camera format of 23-cm square; and end laps of 55, 60, and 65 percent.

7-7. Calculate the approximate vertical exaggeration in a stereomodel from photos taken with a 152-mm-focal-length camera having a 23-cm-square format if the photos are taken at 55 percent end lap assuming 65-mm eye base, 11-cm stereoscope height above the photos and 5.5-cm separation of the photo images.

7-8. Repeat Prob. 7-7, except that a 210-mm-focal-length camera was used, and end lap was 65 percent.

CHAPTER 8

Stereoscopic Parallax

8-1 Introduction

Parallax is the apparent displacement in the position of an object, with respect to a frame of reference, caused by a shift in the position of observation. A simple experiment will serve to illustrate parallax. If a finger is held in front of the eyes, and while gazing at the finger the head is quickly shifted from side to side without moving the finger, the finger will appear to move from side to side with respect to objects beyond the finger, such as pictures on the wall. Rather than shifting the head, the same effect can be created by alternately blinking one's eyes. The closer the finger is held to the eyes, the greater will be its apparent shift. This apparent motion of the finger is parallax, and it is due to the shift in the position of observation.

If a person looked through the viewfinder of an aerial camera as the aircraft moved forward, images of objects would be seen to move across the field of view. This image motion is another example of parallax caused by shifting the location of the observation point. Again, the closer an object is to the camera, the more its image will appear to move.

An aerial camera exposing overlapping photographs at regular intervals of time obtains a record of positions of images at the instants of exposure. The change in position of an image from one photograph to the next caused by the aircraft's motion is termed *stereoscopic parallax*, *x parallax*, or simply *parallax*. Parallax exists for all images appearing on successive overlapping photographs. In Fig. 8-1, for example, images of object points A and B appear on a pair of overlapping vertical aerial photographs which were taken from exposure stations L_1 and L_2. Points A and B are imaged at a and b on the left-hand photograph. Forward motion of the aircraft between exposures, however, caused the images to move laterally across the camera focal plane parallel to the flight line, so that on the right-hand photo they appear at a' and b'. Because point A is higher (closer to the camera) than point B, the movement of image a across the focal plane was greater than the movement of image b; in other words, the parallax of point A is greater than the parallax of point B. This calls attention to two important aspects of stereoscopic parallax: (1) The parallax of any point is directly related to the elevation of the point, and (2) parallax is greater for high points than for low points. Variation of parallax with elevation provides the fundamental basis for determining elevations of points from photographic measurements. In fact, X, Y, and Z ground coordinates can be calculated for points based upon their parallaxes. Equations for doing this are presented in Sec. 8-6.

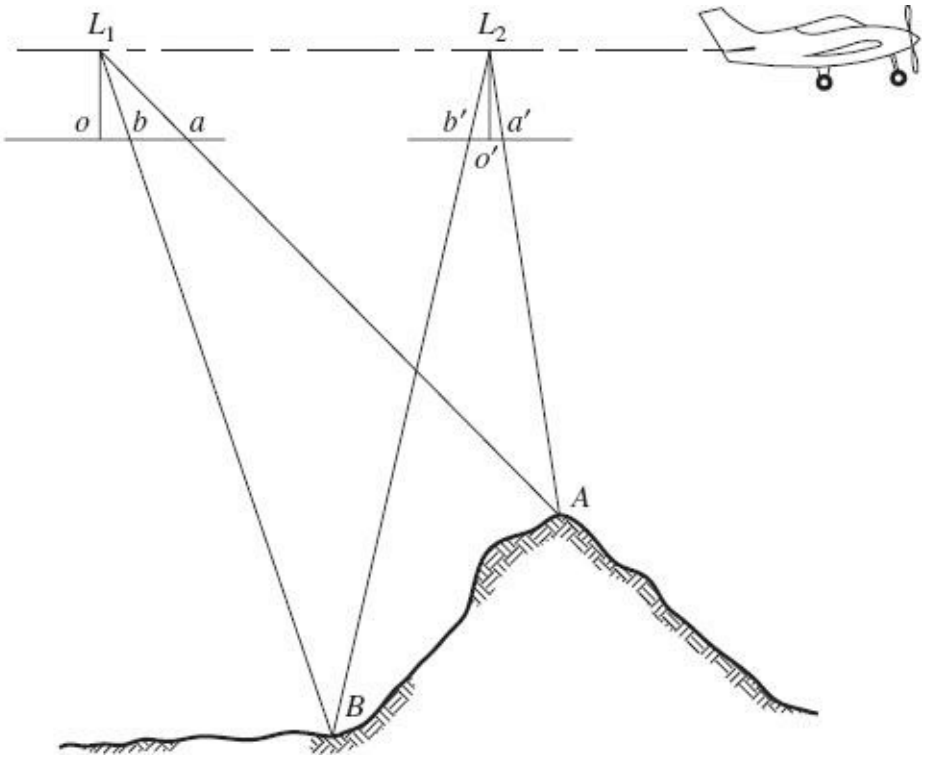


FIGURE 8-1 Stereoscopic parallax of vertical aerial photographs.

Figure 8-2 shows the two photographs of Fig. 8-1 in superposition. Parallaxes of object points A and B are p_a and p_b, respectively. Stereoscopic parallax for any point such as A whose images appear on two photos of a stereopair, expressed in terms of *flight-line* photographic coordinates, is

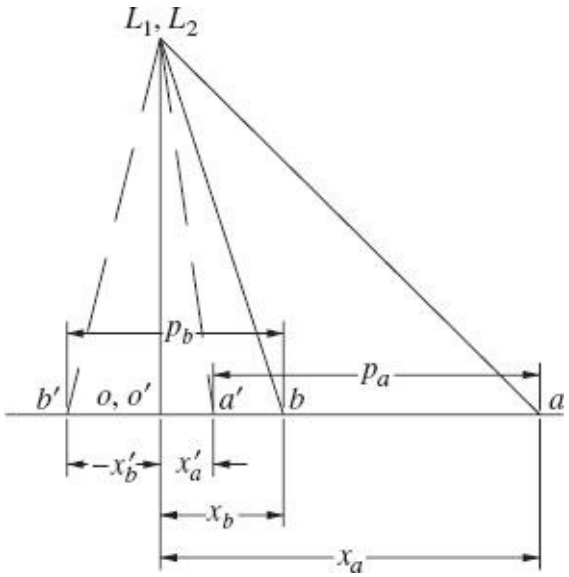


FIGURE 8-2 The two photographs of Fig. 8-1 are shown in superposition.

$$p_a = x_a - x'_a \tag{8-1}$$

In Eq. (8-1), p_a is the stereoscopic parallax of object point A, x_a is the measured photo coordinate of image a on the left photograph of the stereopair, and x_a' is the photo coordinate of image a' on the right photo. These photo coordinates *are not* measured with respect to the fiducial axis system which

was described in Sec. 4-2. Rather, they are measured with respect to the flight-line axis system described in Sec. 8-2. In Eq. (8-1) it is imperative that proper algebraic signs be given to measured photo coordinates to obtain correct values for stereoscopic parallax.

Figure 8-3 is a portion of a stereopair of vertical photographs taken over the University of Florida campus with a 153-mm-focal-length camera at a flying height of 462 m above ground. On these photos, note how all images moved laterally with respect to the *y* axis from their positions on the left photo to their positions on the right photo. Note also how clearly the bell tower (Century Tower) illustrates the increase in parallax with higher points; i.e., the top of the tower has moved farther across the focal plane than the bottom of the tower.

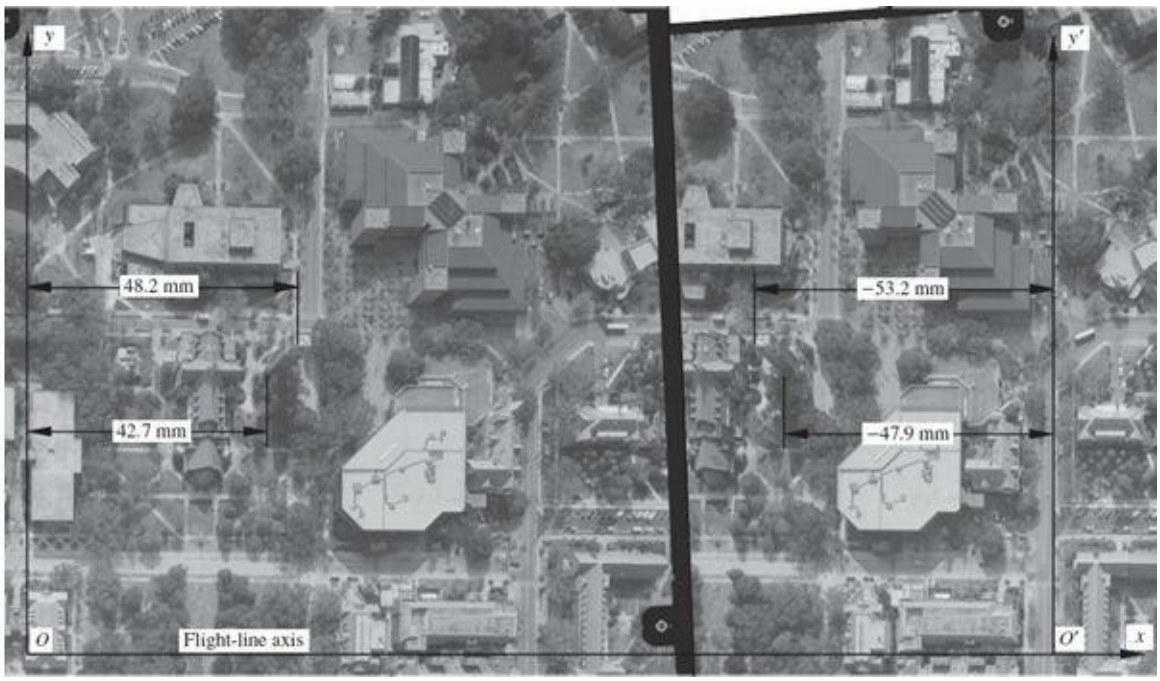


FIGURE 8-3 Overlapping vertical photographs taken over the University of Florida campus illustrating stereoscopic parallax. (*Photos courtesy Hoffman and Company, Inc.*)

In Fig. 8-3, the tower affords an excellent example for demonstrating the use of Eq. (8-1) for finding parallaxes. The top of the tower has an *x* coordinate (x_t = 48.2 mm) and an *x′* coordinate (x_t' = –53.2 mm). By Eq. (8-1), the parallax p_t = 48.2 – (–53.2) = 101.4 mm. Also, the bottom of the tower has an *x* coordinate (x_b = 42.7 mm) and an *x′* coordinate (x_b' = –47.9 mm). Again by Eq. (8-1), p_b = 42.7 – (–47.9) = 90.6 mm.

8-2 Photographic Flight-Line Axes for Parallax Measurement

Since parallax occurs parallel to the direction of flight, the photographic *x* and *x′* axes for parallax measurement must be parallel with the flight line for each of the photographs of a stereopair. (Primed values denote the right-hand photo of a stereopair.) For a vertical photograph of a stereopair, the flight line is the line connecting the principal point and corresponding (conjugate) principal point. The flight line in this context is always a straight line even though the aircraft rarely travels in a perfectly straight line from one camera position to the next. Principal points are located by intersecting the *x* and *y* fiducial lines. A monoscopic method of establishing corresponding principal points is described in Sec. 7-6. Stereoscopic methods are discussed in Sec. 8-4. The *y* and *y′* axes for parallax

measurement pass through their respective principal points and are perpendicular to the flight line.

All photographs except those on the ends of a flight strip may have two sets of flight axes for parallax measurements—one to be used when the photo is the left photo of the stereopair and one when it is the right photo. An example is shown in Fig. 8-4, where photographs 1 through 3 were exposed as shown. Parallax measurements in the overlap area of photos 1 and 2 are made with respect to the solid xy axis system of photo 1 and the solid $x'y'$ system of photo 2. However, due to the aircraft's curved path of travel, the flight line of photos 2 and 3 is not in the same direction as the flight line of photos 1 and 2. Therefore, parallax measurements in the overlap area of photos 2 and 3 must be made with respect to the dashed xy axis system on photo 2 and the dashed $x'y'$ system of photo 3. It is possible for the two axis systems to be coincident; however, this does not generally occur in practice. Henceforth in this chapter it is understood that photographic coordinates for parallax determination are measured with respect to the flight-line axis system.

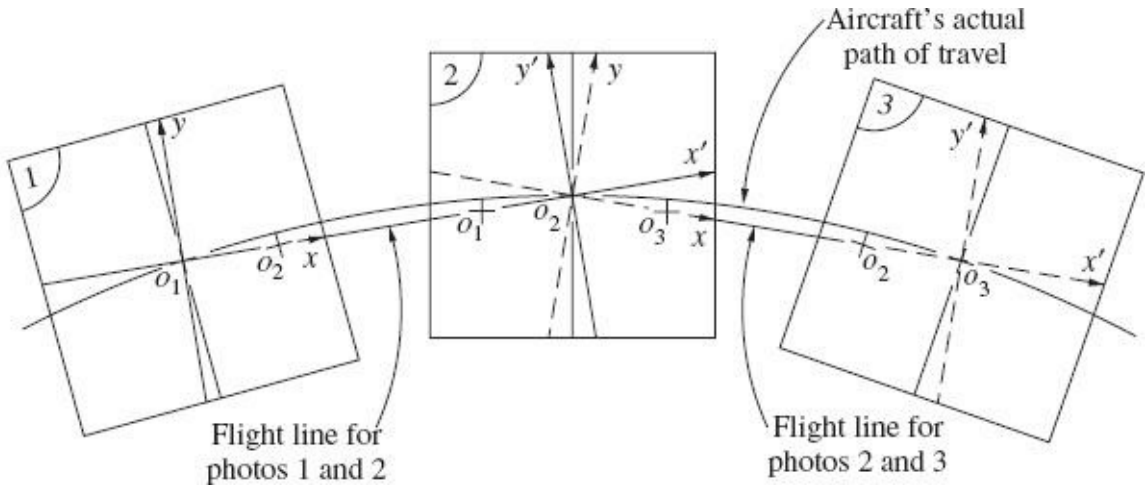


FIGURE 8-4 Flight-line axes for measurement of stereoscopic parallax.

8-3 Monoscopic Methods of Parallax Measurement

Parallaxes of points on a stereopair may be measured either monoscopically or stereoscopically. There are certain advantages and disadvantages associated with each method. In either method the photographic flight line axes must first be carefully located by marking principal points and corresponding principal points.

The simplest method of parallax measurement is the monoscopic approach, in which Eq. (8-1) is solved after direct measurement of x and x' on the left and right photos, respectively. A disadvantage of this method is that two measurements are required for each point.

Another monoscopic approach to parallax measurement is to fasten the photographs down on a table or base material, as shown in Fig. 8-5. In this method the photographic flight lines o_1o_2 and $o_1'o_2'$ are marked as usual. A long straight line AA' is drawn on the base material, and the two photos are *carefully* mounted as shown so that the photographic flight lines are coincident with this line. Now that the photos are fastened down, the distance D between the two principal points is a constant which can be measured. The parallax of point B is $p_b = x_b - x_b'$ (note that in Fig. 8-5 the x_b' coordinate is negative). However, by examining the figure, it is seen that parallax is also

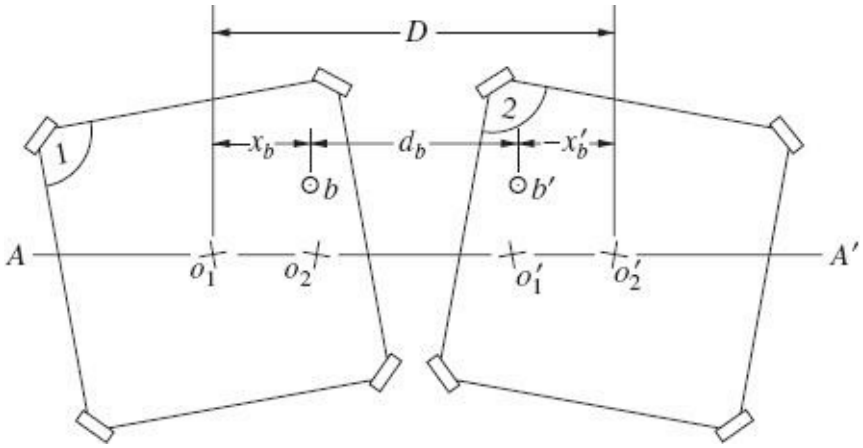


FIGURE 8-5 Parallax measurement using a simple scale.

$$p_b = D - d_b \tag{8-2}$$

With D known in Eq. (8-2), to obtain the parallax of a point it is necessary only to measure the distance d between its images on the left and right photos. The advantage is that for each additional point whose parallax is desired, only a single measurement is required. With either of these monoscopic methods of parallax measurement, a simple scale as described in Sec. 4-3 may be used, with the choice being based upon the desired accuracy.

8-4 Principle of the Floating Mark

Parallaxes of points can be measured while viewing stereoscopically with the advantages of speed and accuracy. Stereoscopic measurement of parallax makes use of the principle of the *floating mark*. When a stereomodel is viewed through a stereoscope, two small identical marks printed on clear plastic transparencies, called *half marks*, may be placed over the photographs—one on the left photo and one on the right photo, as illustrated in Fig. 8-6. The left mark is seen with the left eye and the right mark with the right eye. The half marks may be shifted in position until they fuse together into a single mark which appears to exist in the stereomodel and to lie at a particular elevation. If the half marks are moved closer together, the parallax of the half marks is increased and the fused mark will therefore appear to rise. Conversely, if the half marks are moved apart, parallax is decreased and the fused mark appears to fall. This apparent variation in the elevation of the mark as the spacing of half marks is varied is the basis for the term *floating mark*. Usable half marks can be created by printing two plus signs (+) several centimeters apart on a plastic transparency. Pinholes are then precisely made at the center of the crosses. Finally two small pieces approximately 2-cm square can be cut from the plastic with the crosses at the centers.

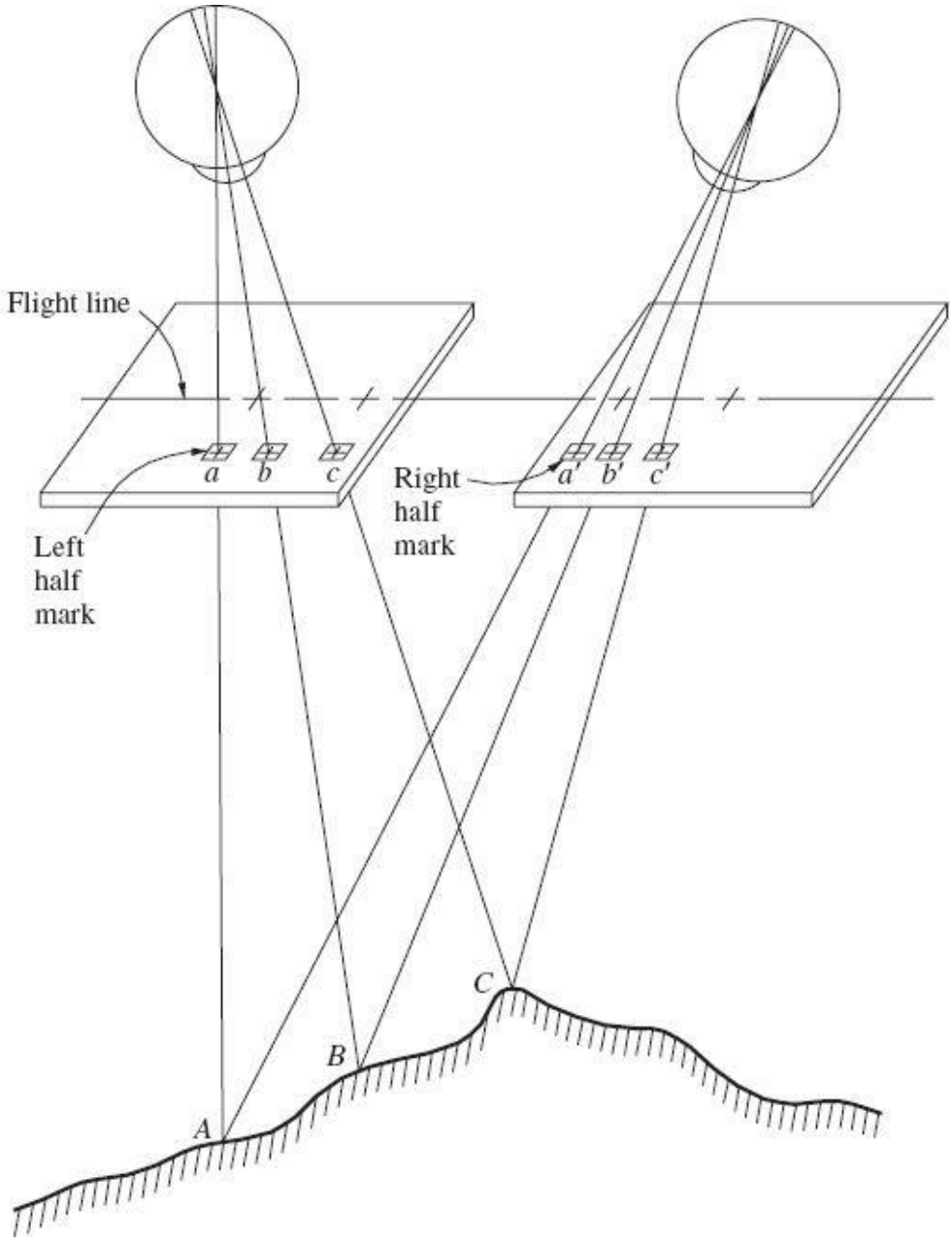


FIGURE 8-6 The principle of the floating mark.

The spacing of the half marks, and hence the parallax of the half marks, may be varied so that the floating mark appears to rest exactly on the terrain. This produces the same effect as though an object of the shape of the half marks had existed on the terrain when the photos were originally taken. The floating mark may be moved about the stereomodel from point to point, and as the terrain varies in elevation, the spacing of the half marks may be varied to make the floating mark rest exactly on the terrain. Figure 8-6 demonstrates the principle of the floating mark and illustrates how the mark may be set exactly on particular points such as *A*, *B*, and *C* by placing the half marks at *a* and *a'*, *b* and *b'*, and *c* and *c'*, respectively.

The principle of the floating mark can be used to transfer principal points to their corresponding locations, thereby marking the flight-line axes. In this procedure the principal points are first located as usual at the intersection of fiducial lines. Then the left half mark is placed over one of the principal points, say, the left point o_1. It is held fixed in that position. Using a mirror stereoscope for viewing, the right half mark is placed on the right photo and moved about until a clear stereoscopic view of the floating mark is obtained and the fused mark appears to rest exactly on the ground. The right half mark is then carefully held in place while a pin is inserted through the center of the cross to make a

pinprick on the photograph. This stereoscopic procedure is very accurate if carefully performed, and it has the advantage that discrete images near the principal points are not necessary, as they are with the monoscopic method. Imagine, e.g., the difficulty of monoscopically transferring a principal point that falls in the middle of a grassland. This transfer could be readily done by the stereoscopic method, however.

Once corresponding principal points have been marked, the *photo base b* can be determined. The photo base is the distance on a photo between the principal point and the corresponding principal point from the overlapping photo. Figure 8-7 is a vertical section through the exposure stations of a pair of overlapping vertical photos. By Eq. (8-1), the parallax of the left-photo ground principal point P_1 is $p_{o1} = x_{o1} - (-x'_{o1}) = 0 - (-b') = b'$. (The x coordinate of o_1 on the left photo is zero.) Also, the parallax of the right-photo ground principal point P_2 is $p_{o2} = x_{o2} - (-x'_{o2}) = b - 0 = b$. From the foregoing, it is seen that *the parallax of the left ground principal point is photo base b′ measured on the right photo*, and *the parallax of the right ground principal point is photo base b measured on the left photo*. In areas of moderate relief, the values of b and b' will be approximately equal, and the photo base for the stereopair can be taken as the average of these two values.

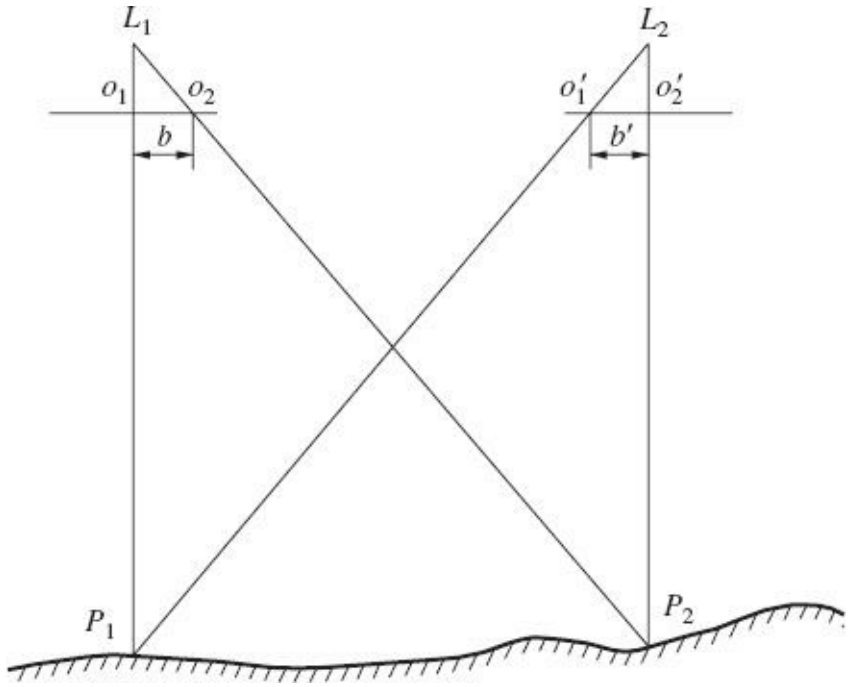


FIGURE 8-7 Parallax of the principal points.

8-5 Stereoscopic Methods of Parallax Measurement

Through the principle of the floating mark, parallaxes of points may be measured stereoscopically. This method employs a stereoscope in conjunction with an instrument called a *parallax bar*, also frequently called a *stereometer*. A parallax bar consists of a metal rod to which are fastened two half marks. The right half mark may be moved with respect to the left mark by turning a micrometer screw. Readings from the micrometer are taken with the floating mark set exactly on points whose parallaxes are desired. From the micrometer readings, parallaxes or differences in parallax are obtained. A parallax bar is shown lying on the photos beneath a mirror stereoscope in Fig. 8-8.

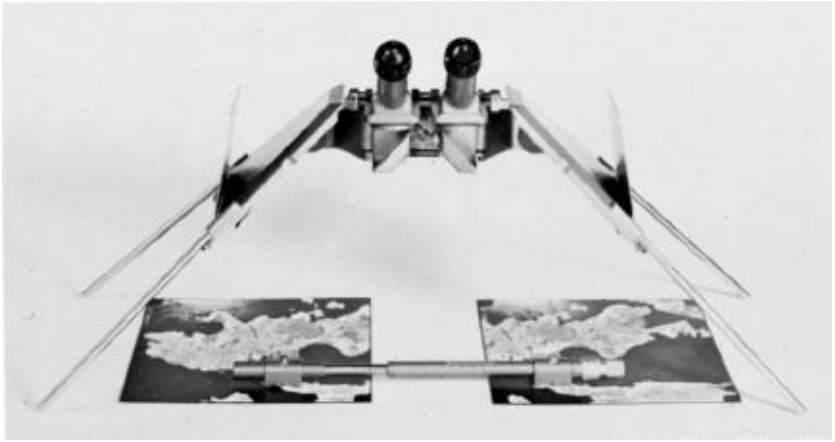


FIGURE 8-8 Wild ST-4 mirror stereoscope with binocular attachment and parallax bar. (*Courtesy LH Systems, LLC.*)

When a parallax bar is used, the two photos of a stereopair are first *carefully* oriented for comfortable stereoscopic viewing, in such a way that the flight line of each photo lies precisely along a common straight line, as line AA' shown in Fig. 8-5. The photos are then fastened securely, and the parallax bar is placed on the photos. The left half mark, called the *fixed mark*, is unclamped and moved so that when the floating mark is fused on a terrain point of average elevation, the parallax bar reading is approximately in the middle of the run of the graduations. The fixed mark is then clamped, where it will remain for all subsequent parallax measurements on that particular stereopair. After the fixed mark is positioned in this manner, the right half mark, or *movable mark*, may be moved left or right with respect to the fixed mark (increasing or decreasing the parallax) as required to accommodate high points or low points without exceeding the run of the parallax bar graduations.

Figure 8-9 is a schematic diagram illustrating the operating principle of the parallax bar. After the photos have been oriented and the left half mark is fixed in position as just described, the *parallax bar constant* C for the setup is determined. For the setup, the spacing between principal points is a constant, denoted by D. Once the fixed mark is clamped, the distance from the fixed mark to the index mark of the parallax bar is also a constant, denoted by K. From Fig. 8-9, the parallax of point A is

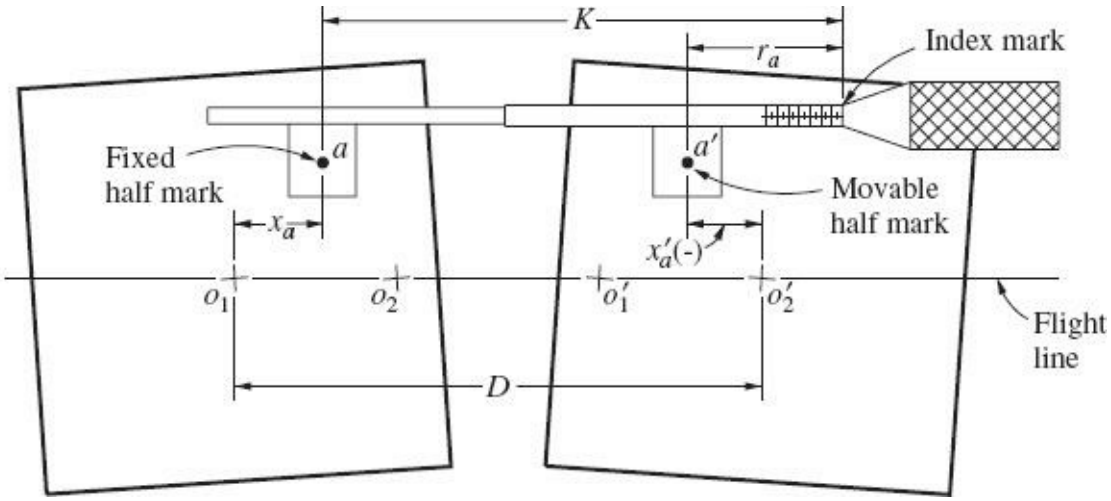


FIGURE 8-9 Schematic diagram of the parallax bar.

$$p_a = x_a - x'_a = D - (K - r_a) = (D - K) + r_a$$

The term ($D - K$) is C, the parallax bar constant for the setup. Also r_a is the micrometer reading. By substituting C into the above equation, the expression becomes

$$p_a = C + r_a \tag{8-3}$$

To calculate the parallax bar constant, a micrometer reading is taken with the floating mark set on a selected point. The parallax of that point is also directly measured monoscopically and calculated using Eq. (8-1). Then with p and r for that point known, the value of C is calculated by using Eq. (8-3), as

$$C = p - r \tag{8-4}$$

The parallax bar constant should be determined on the basis of micrometer readings and parallax measurements for two points. Then the mean of the two values may be adopted. Any two points may be selected for this purpose; however they should be clear, discrete images, and selected so that they lie on opposite sides of the flight line and approximately equidistant from the flight line. This minimizes error in parallaxes due to tilt and faulty orientation of the photos.

One of the advantages of measuring parallax stereoscopically is increased speed, for once the parallax bar constant is determined, the parallaxes of all other points are quickly obtained with a single micrometer reading for each point. Another advantage is increased accuracy. An experienced person using quality equipment and clear photos is generally able to obtain parallaxes to within approximately 0.03 mm of their correct values.

8-6 Parallax Equations

As noted earlier, X, Y, and Z ground coordinates can be calculated for points based upon the measurements of their parallaxes. Figure 8-10 illustrates an overlapping pair of vertical photographs which have been exposed at equal flying heights above datum. Images of an object point A appear on the left and right photos at a and a', respectively. The planimetric position of point A on the ground is given in terms of ground coordinates X_A and Y_A. Its elevation above datum is h_A. The XY ground axis system has its origin at the datum principal point P of the left-hand photograph; the X axis is in the same vertical plane as the photographic x and x' flight axes; and the Y axis passes through the datum principal point of the left photo and is perpendicular to the X axis. According to this definition, each stereopair of photographs has its own unique ground coordinate system.

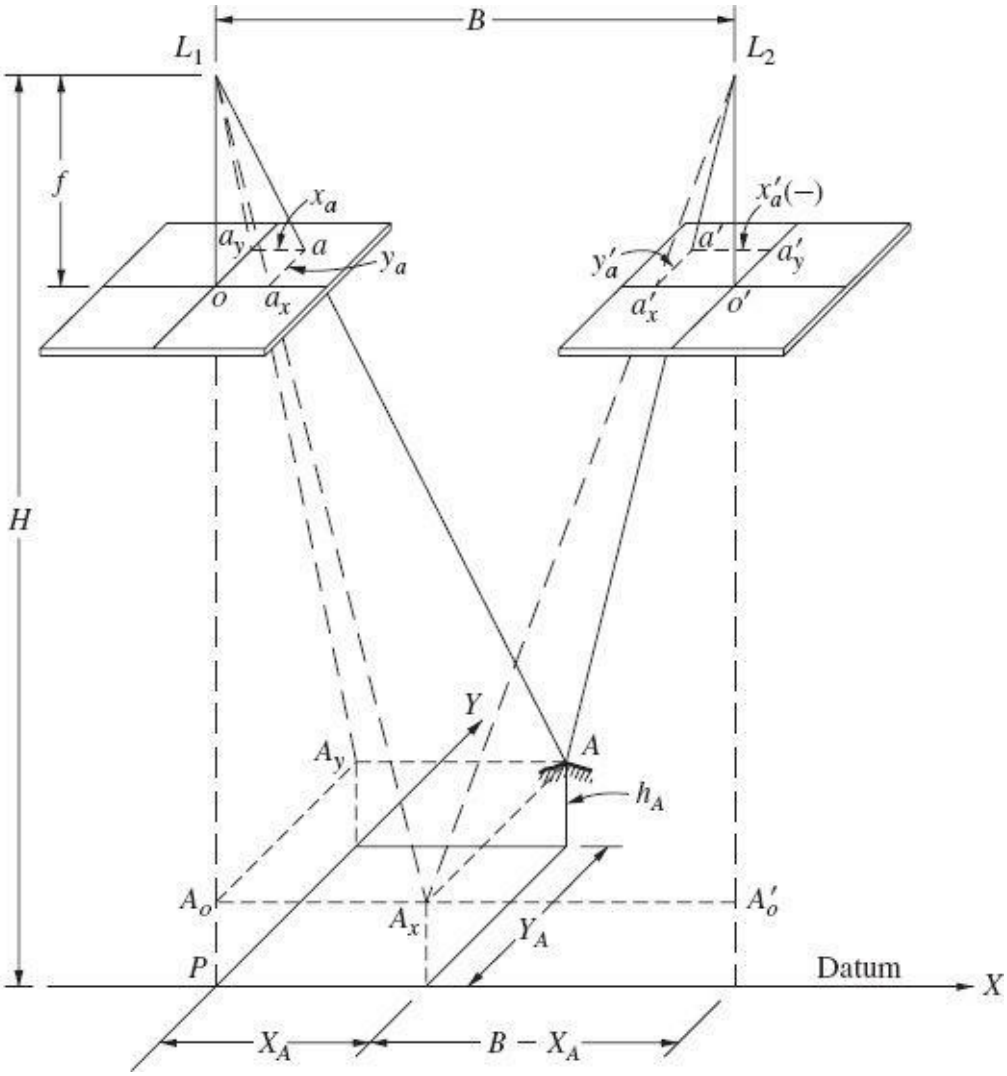


FIGURE 8-10 Geometry of an overlapping pair of vertical photographs.

By equating similar triangles of Fig. 8-10, formulas for calculating h_A, X_A, and Y_A may be derived. From similar triangles L_1oa_y and $L_1A_oA_y$,

$$\frac{Y_A}{H-h_A}=\frac{y_a}{f}$$

from which

$$Y_A=\frac{y_a}{f}(H-h_A) \qquad (a)$$

and equating similar triangles L_1oa_x and $L_1A_oA_x$, we have

$$\frac{X_A}{H-h_A}=\frac{x_a}{f}$$

from which

$$X_A=\frac{x_a}{f}(H-h_A)$$

(*b*)

Also from similar triangles $L_2\, o'a_x'$ and $L_2\, A_o'\, A_x$,

$$\frac{B - X_A}{H - h_A} = \frac{-x_a'}{f}$$

from which

$$X_A = B + \frac{x_a'}{f}(H - h_A)$$

(*c*)

Equating Eqs. (*b*) and (*c*) and reducing gives

$$h_A = H - \frac{Bf}{x_a - x_a'}$$

(*d*)

Substituting p_a for $x_a - x_a'$ into Eq. (*d*) yields

$$h_A = H - \frac{Bf}{p_a}$$

(8-5)

Now substituting Eq. (8-5) into each of Eqs. (*b*) and (*a*) and reducing gives

$$X_A = B\frac{x_a}{p_a}$$

(8-6)

$$Y_A = B\frac{y_a}{p_a}$$

(8-7)

In Eqs. (8-5), (8-6), and (8-7), h_A is the elevation of point *A* above datum, *H* is the flying height above datum, *B* is the air base, *f* is the focal length of the camera, p_a is the parallax of point *A*, X_A and Y_A are the ground coordinates of point *A* in the previously defined unique arbitrary coordinate system, and x_a and y_a the photo coordinates of point *a* measured with respect to the flight-line axes on the left photo.

Equations (8-5), (8-6), and (8-7) are commonly called the *parallax equations*. These equations enable a moderate accuracy survey of the overlap area of a stereopair to be made, provided the focal length is known and sufficient ground control is available so the air base *B* and flying height *H* can be calculated.

Equations (8-6) and (8-7) yield *X* and *Y* ground coordinates in the unique arbitrary coordinate system of the stereopair, which is not related to any standard two-dimensional ground coordinate system. However, if arbitrary *XY* coordinates are determined using these equations for at least two points whose ground coordinates are also known in a standard two-dimensional coordinate system (e.g., state plane coordinates), then the arbitrary *XY* coordinates of all other points can be transformed into that ground system through a two-dimensional coordinate transformation, as described in App. C.

Example 8-1
A pair of overlapping vertical photographs was taken from a flying height of 1233 m above sea level with a 152.4-mm-focal-length camera. The air base was 390 m. With the photos properly oriented, flight-line coordinates for points *a* and *b* were measured as x_a = 53.4 mm, y_a = 50.8 mm, x_a' = –38.3 mm, y_a' = 50.9 mm, x_b = 88.9 mm, y_b = –46.7 mm, x_b' = –7.1 mm, y_b' = –46.7 mm. Calculate the elevations of points *A* and *B* and the horizontal length of line *AB*.

Solution By Eq. (8-1)

$$p_a = x_a - x_a' = 53.4 - (-38.3) = 91.7 \text{ mm}$$

$$p_b = x_b - x_b' = 88.9 - (-7.1) = 96.0 \text{ mm}$$

By Eq. (8-5),

$$h_A = H\frac{Bf}{p_a} = 1233 - \frac{390(152.4)}{91.7} = 585 \text{ m above sea level}$$

$$h_B = H\frac{Bf}{p_b} = 1233 - \frac{390(152.4)}{96.0} = 614 \text{ m above sea level}$$

By Eqs. (8-6) and (8-7),

$$X_A = B\frac{x_a}{p_a} = 390\left(\frac{53.4}{91.7}\right) = 227 \text{ m}$$

$$Y_A = B\frac{y_a}{p_a} = 390\left(\frac{50.8}{91.7}\right) = 216 \text{ m}$$

$$X_B = B\frac{x_b}{p_b} = 390\left(\frac{88.9}{96.0}\right) = 316 \text{ m}$$

$$Y_B = B\frac{y_b}{p_b} = 390\left(\frac{-46.7}{96.0}\right) = -190 \text{ m}$$

The horizontal length of line *AB* is

$$AB = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2} = \sqrt{(361-227)^2 + (-190-216)^2} = 427 \text{ m}$$

▲

8-7 Elevations by Parallax Differences

Parallax differences between one point and another are caused by different elevations of the two points. While parallax Eq. (8-5) serves to define the relationship of stereoscopic parallax to flying height, elevation, air base, and camera focal length, parallax differences are more convenient for

determining elevations. In Fig. 8-11, object point C is a control point whose elevation h_C above datum is known. The elevation of object point A is desired. By rearranging Eq. (8-5), parallaxes of both points can be expressed as

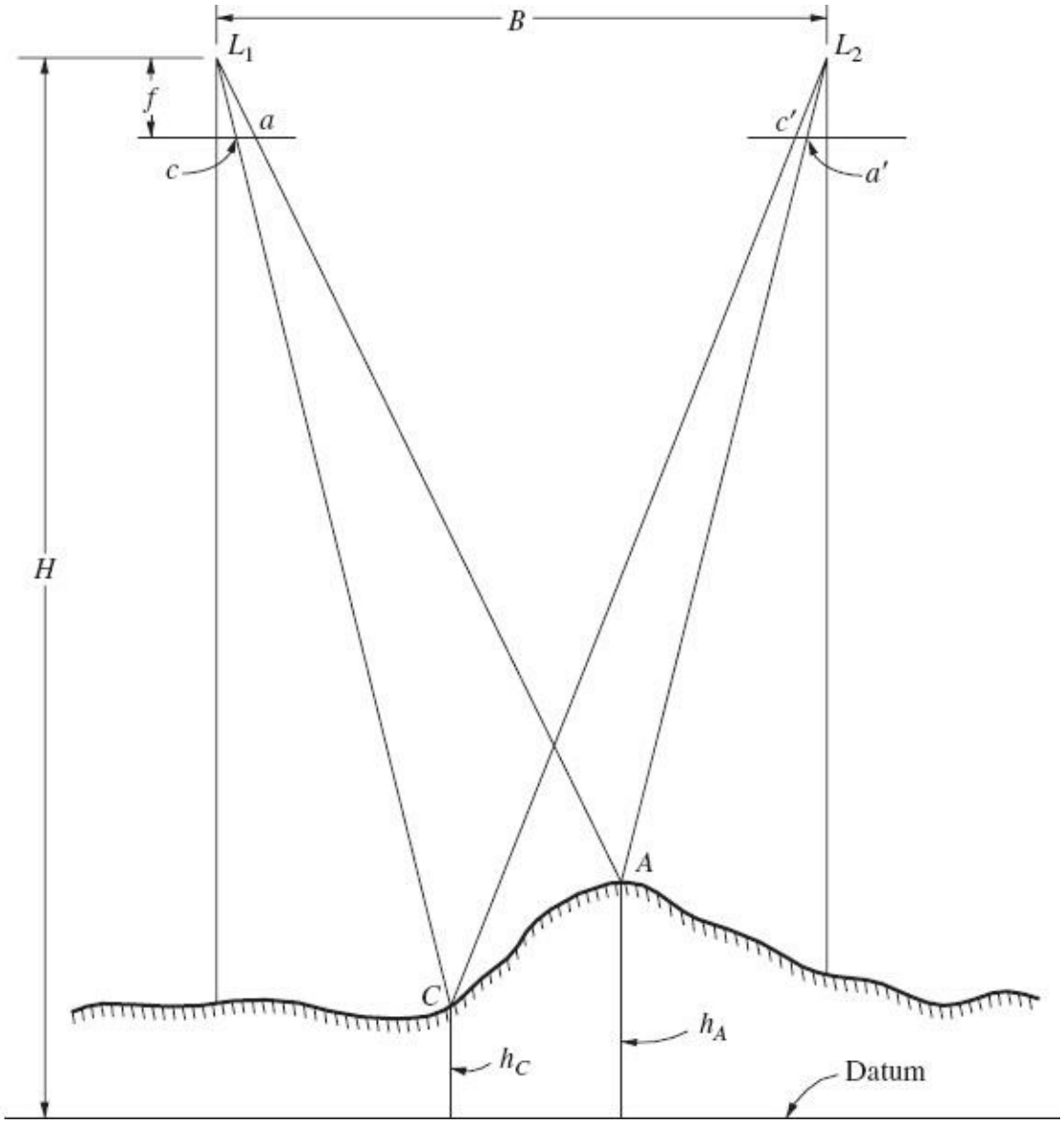


FIGURE 8-11 Elevations by parallax differences.

$$p_c = \frac{fB}{H - h_C} \tag{e}$$

$$p_a = \frac{fB}{H - h_A} \tag{f}$$

The difference in parallax $p_a - p_c$, obtained by subtracting Eq. (e) from Eq. (f) and rearranging, is

$$p_a - p_c = \frac{fB(h_A - h_C)}{(H - h_A)(H - h_C)} \tag{g}$$

Let $p_a - p_c$ equal Δp, the difference in parallax. By substituting $H - h_A$ from Eq. (*f*), and Δp into (*g*) and reducing, the following expression for elevation h_A is obtained:

$$h_A = h_C + \frac{\Delta p(H - h_C)}{p_a} \tag{8-8}$$

Example 8-2
In Example 8-1, flight-line axis x and x' coordinates for the images of a vertical control point C were measured as x_c = 14.3 mm and x_c' = –78.3 mm. If the elevation of point C is 591 m above sea level, calculate the elevations of points A and B of that example, using parallax difference Eq. (8-8).

Solution By Eq. (8-1),

$$p_c = x_c - x_c' = 14.3 - (-78.3) = 92.6 \text{ mm}$$

For point *A*,

$$\Delta p = p_a - p_c = 91.7 - 92.6 = -0.9 \text{ mm}$$

By Eq. (8-8),

$$h_A = h_C + \frac{\Delta p(H - h_C)}{p_a} = 591 + \frac{(-0.9)(233 - 591)}{91.7}$$

$$= 585 \text{ m above sea level}$$

For point *B*,

$$\Delta p = p_b - p_c = 96.0 - 92.6 = 3.4 \text{ mm}$$

By Eq. (8-8),

$$h_B = h_C + \frac{\Delta p(H - h_C)}{p_b} = 591 + \frac{(3.4)(1233 - 591)}{96.0} = 614 \text{ m above sea level}$$

▲

Note that these answers check the values computed in Example 8-1.

If a number of control points are located throughout the overlap area, use of Eq. (8-8) permits elevations of unknown points to be most accurately determined from the parallax difference of the nearest control point. This minimizes the effects of two primary errors—photographic tilt and imperfect alignment of the photos for parallax measurement.

8-8 Simplified Equation for Heights of Objects from Parallax Differences

In many applications it is necessary to estimate heights of objects to a moderate level of accuracy.

Utilizing parallax differences for height determination is particularly useful when application of relief displacement is not possible because either the feature is not vertical (e.g., a construction crane) or the base of the feature is obscured (e.g., trees in a forest). In a situation like this, a parallax difference can be determined between a point on the ground and the top of the feature. A fundamental assumption is, of course, that the point on the ground is at the same elevation as the base of the feature. In a large number of cases, this assumption is valid as long as only moderate accuracy is required.

A simplified equation for height determination can be obtained from Eq. (8-8) by choosing the vertical datum to be the elevation of the point on the ground that is used as the basis for the parallax difference. This makes h_C zero, and Eq. (8-8) simplifies to

$$h_A = \frac{\Delta p H}{p_a} \tag{8-9}$$

In Eq. (8-9), h_A is the height of point A above ground, $\Delta p = p_a - p_c$ is the difference in parallax between the top of the feature and the ground (p_c is the parallax of the ground), and H is the flying height above ground, since datum is at ground. If the heights of many features are needed in an area where the ground is approximately level, the photo base b can be utilized as the parallax of the ground point. In this case, Eq. (8-9) can be modified to

$$h_A = \frac{\Delta p H}{b + \Delta p} \tag{8-10}$$

In Eq. (8-10), b is the photo base for the stereopair, $\Delta p = p_a - b$, and the other terms are as previously defined. For very low flying heights or in areas of significant relief, or both, the assumptions of Eq. (8-10) are not met; in these cases, Eq. (8-8) should be used. Equation (8-10) is especially convenient in photo interpretation where rough elevations, building and tree heights, etc. are often needed.

Example 8-3
The parallax difference between the top and bottom of a tree is measured as 1.3 mm on a stereopair of photos taken at 915 m above ground. Average photo base is 88.2 mm. How tall is the tree?

Solution By Eq. (8-10),

$$h = \frac{1.3 \times 915}{88.2 + 1.3} = 13 \text{ m} \quad \blacktriangle$$

8-9 Measurement of Parallax Differences

Parallax differences may be determined in any of the following ways:

1. By monoscopic measurement of parallaxes followed by subtraction
2. By taking differences in parallax bar readings
3. By *parallax wedge*

A parallax wedge, as illustrated in Fig. 8-12, consists of a piece of transparent film upon which are drawn two converging lines. The left line is a reference line while the line on the right contains graduations from which readings can be made. The spacing of the two lines depends on whether the parallax wedge will be used with a mirror stereoscope or a pocket stereoscope. For a pocket stereoscope the spacing should vary from about 65 mm at the bottom to about 45 mm at the top. This spacing accommodates the usual spacing between corresponding images when a stereopair is oriented for viewing with a pocket stereoscope, and it gives a possible range of about 20 mm in parallax differences that can be measured.

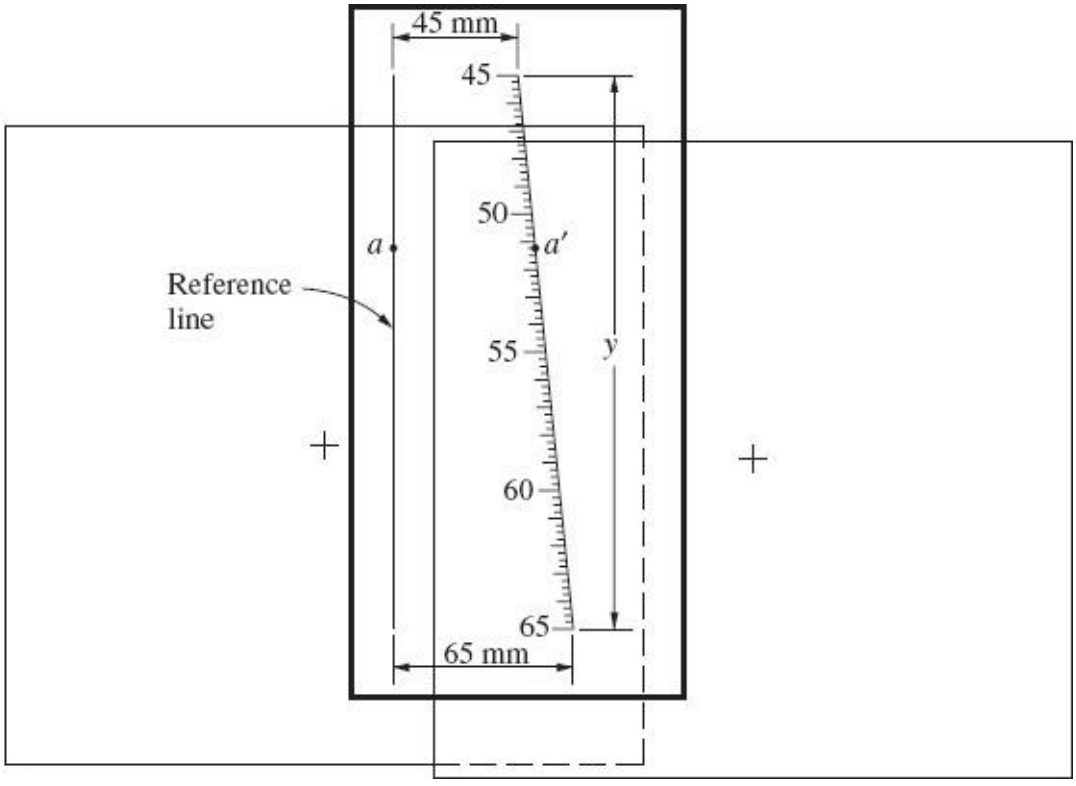


FIGURE 8-12 Parallax wedge.

Suppose line spacings of a parallax wedge were exactly 65 mm at the bottom and 45 mm at the top, as shown in Fig. 8-12. If the total height *y* of the graduations was 200 mm, then for graduations spaced at 2.5-mm intervals along the line, each of the 80 graduations proceeding upward on the scale is 0.25 mm closer to the reference line than the next-lower graduation. Graduations numbered from 45 to 65 on the parallax wedge represent horizontal spacings from the reference line, with the smallest graduation interval representing a parallax difference of 0.25 mm even though the vertical spacing between consecutive lines is 2.5 mm.

When a parallax wedge is used, the photos are first carefully oriented for viewing with a pocket stereoscope and secured. The parallax wedge is placed in the overlap area and viewed stereoscopically; the two lines of the parallax wedge will fuse and appear as a single *floating line* in areas where the spacing of the lines is slightly less than the spacing of corresponding photo images. The floating line will appear to split where the parallax of the lines is equal to that of the photo images. The position of the parallax wedge can be adjusted so that the floating line splits forming a wedge exactly at a point whose parallax is desired, and at that point a reading is taken from the scale. The parallax wedge reading at point *a* of Fig. 8-12, for example, is 51.25 mm. Parallax differences are obtained by simply taking differences in parallax wedge readings for different points.

An expedient means of producing a parallax wedge is to create a drawing using an available computer-aided drafting (CAD) program and to print the result on a transparency using a laser printer.

8-10 Computing Flying Height and Air Base

To use parallax equations, it is generally necessary to compute the flying height and air base. Flying height may be calculated using the methods described in Sec. 6-9. For best results, the average of flying heights for the two photos of a stereopair should be used.

If the air base is known and if one vertical control point is available in the overlap area, flying height for the stereopair may be calculated by using Eq. (8-5).

Example 8-4

An overlapping pair of vertical photographs taken with a 152.4-mm-focal-length camera has an air base of 548 m. The elevation of control point *A* is 283 m above sea level, and the parallax of point *A* is 92.4 mm. What is the flying height above sea level for this stereopair?

Solution By rearranging Eq. (8-5),

$$H = h + \frac{Bf}{p} = 283 + \frac{548(152.4)}{92.4} = 1187 \text{ m above sea level} \quad \blacktriangle$$

If the flying height above datum is known and if one vertical control point is available in the overlap area, the air base for the stereopair may be calculated by using Eq. (8-5).

Example 8-5

An overlapping pair of vertical photos was exposed with a 152.4-mm-focal-length camera from a flying height of 1622 m above datum. Control point *C* has an elevation of 263 m above datum, and the parallax of its images on the stereopair is 86.3 mm. Calculate the air base.

Solution By rearranging Eq. (8-5),

$$B = (H - h)\frac{p}{f} = (1622 - 263)\frac{86.3}{152.4} = 770 \text{ m}$$

If a line of known horizontal length appears in the overlap area, then the air base can be readily calculated. The horizontal length of the line may be expressed in terms of rectangular coordinates, according to the pythagorean theorem, as

$$AB = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2}$$

Substituting Eqs. (8-6) and (8-7) into the above for the rectangular coordinates gives

$$AB = \sqrt{\left(\frac{Bx_b}{p_b} - \frac{Bx_a}{p_a}\right)^2 + \left(\frac{By_b}{p_b} - \frac{By_a}{p_a}\right)^2}$$

Solving the above equation for *B* yields

$$B = \frac{AB}{\sqrt{(x_b/p_b - x_a/p_a)^2 + (y_b/p_b - y_a/p_a)^2}} \tag{8-11}$$

▲

Example 8-6
Images of the endpoints of ground line *AB*, whose horizontal length is 650.47 m, appear on a pair of overlapping vertical photographs. Photo coordinates measured with respect to the flight axis on the left photo were x_a = 33.3 mm, y_a = 13.5 mm, x_b = 41.8 mm, and y_b = –95.8 mm. Photo coordinates measured on the right photo were x_a' = –52.3 mm and x_b' = –44.9 mm. Calculate the air base for this stereopair.

Solution By Eq. (8-1),

$$p_a = x_a - x_a' = 33.3 - (-52.3) = 85.6 \text{ mm}$$
$$p_b = x_b - x_b' = 41.8 - (-44.9) = 86.7 \text{ mm}$$

By Eq. (8-9),

$$B = \frac{650.47}{\sqrt{(41.8/86.7 - 33.3/85.6)^2 + (-95.8/86.7 - 13.5/85.6)^2}} = 514 \text{ m}$$

▲

8-11 Error Evaluation

Answers obtained using the various equations presented in the chapter will inevitably contain errors. It is important to be aware of the presence of these errors and to be able to assess their magnitudes. Some of the sources of error in computed answers using parallax equations are as follows:

1. Locating and marking the flight lines on photos
2. Orienting stereopairs for parallax measurement
3. Parallax and photo coordinate measurements
4. Shrinkage or expansion of photographs
5. Unequal flying heights for the two photos of stereopairs
6. Tilted photographs
7. Errors in ground control
8. Other errors of lesser consequence such as camera lens distortion and atmospheric refraction distortion

A general approach for determining the combined effect of several random errors in computed answers is presented in Sec. A-4. This approach is demonstrated in the following example.

Example 8-7
In the computation of the elevation of point *A* in Example 8-1, suppose that the random errors were ± 2 m in *H*, ± 2 m in *B*, and ± 0.1 mm in p_a. Compute the resulting error in h_A due to the presence of

these errors.

Solution The basic equation used was Eq. (8-5), and the partial derivatives in that equation taken with respect to each of the three error sources are

$$\frac{\partial h_A}{\partial H} = 1 \qquad \frac{\partial h_A}{\partial B} = -\frac{f}{p_a} \qquad \frac{\partial h_A}{\partial p_a} = \frac{Bf}{p_a^2}$$

Substituting the above expressions into Eq. (A-2) gives

$$\sigma_{h_A} = \pm \sqrt{(1)^2 \sigma_H^2 + \left(\frac{-f}{p_a}\right)^2 \sigma_B^2 + \left(\frac{Bf}{p_a^2}\right)^2 \sigma_{p_a}^2}$$

Substituting numerical values into the above, we get

$$\sigma_{h_A} = \pm \sqrt{(1)^2(2)^2 + \left(\frac{-152.4}{91.7}\right)^2 (2)^2 + \left[\frac{390(152.4)}{(91.7)^2}\right]^2 (0.1)^2}$$

$$= \pm \sqrt{4 + 11 + 0.5} = \pm 3.9 \text{ m} \quad \blacktriangle$$

Errors in computed answers using any of the equations of this chapter can be analyzed in the fashion described above. It is, of course, necessary to estimate the magnitude of the random errors in the measured variables used in the equations. It is more difficult to analyze errors caused by tilt in the photographs. The subject of tilted photographs is discussed in Chap. 10, and rigorous analytical methods are presented in Chap. 11. For the present, however, suffice it to say that for normal photography intended to be vertical, errors in parallax equation answers due to tilt are compatible with errors from the other sources that have been considered.

References

American Society of Photogrammetry: *Manual of Photogrammetry*, 3d ed., Bethesda, MD, 1966, chap. 2.

———: *Manual of Photogrammetry*, 4th ed., Bethesda, MD, 1980, chap. 2.

Avery, T. E.: "Two Cameras for Parallax Height Measurements," *Photogrammetric Engineering*, vol. 32, no. 6, 1966, p. 576.

Bender, L. U.: "Derivation of Parallax Equation," *Photogrammetric Engineering*, vol. 33, no. 10, 1967, p. 1175.

Nash, A. J.: "Use a Mirror Stereoscope Correctly," *Photogrammetric Engineering*, vol. 38, no. 12, 1972, p. 1192.

Porter, G. R.: "Errors in Parallax Measurements and Their Assessment in Student Exercises," *Photogrammetric Record*, vol. 8, no. 46, 1975, p. 528.

Schut, G. H.: "The Determination of Tree Heights from Parallax Measurements," *Canadian Surveyor*, vol. 19, 1965, p. 415.

Problems

8-1. Calculate the stereoscopic parallaxes of points *A* through *D*, given the following measured flight-line axis coordinates. Which point is the highest in elevation? Which is lowest?

Point	*x* (Left Photo)	*x*′ (Right Photo)
A	21.3 mm	−67.2 mm
B	53.8 mm	−33.7 mm
C	71.5 mm	−20.1 mm
D	103.0 mm	12.9 mm

8-2. Calculate the elevations of points *A* through *D* of Prob. 8-1 if the camera focal length is 152.62 mm, flying height above datum is 2763 m, and the air base is 1135 m.

8-3. A pair of overlapping vertical photographs is mounted for parallax measurement, as illustrated in Fig. 8-5. Distance *D* is measured as 276.1 mm. Calculate the stereoscopic parallaxes of the points whose measured *d* values are as follows. Which point is highest in elevation? Which is lowest?

Point	***d***
A	183.0 mm
B	185.5 mm
C	186.4 mm
D	187.8 mm

8-4. Repeat Prob. 8-3, except *D* was measured as 272.7 mm, and measured *d* values are as follows.

Point	***d***
A	182.5 mm
B	184.7 mm
C	185.3 mm
D	180.2 mm

8-5. Assume that point *A* of Prob. 8-3 has an elevation of 357.1 m above datum and that the photos were taken with an 88.74-mm-focal-length camera. If the air base is 1357 m, what are the elevations of points *B*, *C*, and *D*?

8-6. Assume that point *A* of Prob. 8-4 has an elevation of 249.0 m above datum and that the photos were taken with a 153.07-mm-focal-length camera. If the air base is 1773 m, what are the elevations of points *B*, *C*, and *D*?

8-7. From the information given for Probs. 8-1 and 8-2, calculate the horizontal ground length of line *AC*. Measured *y* coordinates on the left photo are y_a = 70.8 mm and y_c = –64.9 mm.

8-8. Repeat Prob. 8-7 except that the computations are for line *BD*. Measured *y* coordinates on the left photo are y_b = 10.8 mm and y_d = –81.5 mm.

8-9. From the data of Probs. 8-3 and 8-5, calculate the horizontal area of triangle *ABC*. Measured *x* and *y* flight-line axis coordinates of *a*, *b*, and *c* on the left photo were $x_a = -7.6$ mm, $y_a = 90.3$ mm, $x_b = 31.8$ mm, $y_b = -79.3$ mm, $x_c = 88.4$ mm, and $y_c = 16.1$ mm.

8-10. Distances *b* on the left photo and *b′* on the right photo of a pair of overlapping vertical photos are 91.0 mm and 88.2 mm, respectively. If the air base is 604 m and the camera focal length is 88.78 mm, which ground principal point is higher and by how much?

8-11. Repeat Prob. 8-10, except that *b* and *b′* are 90.6 mm and 92.1 mm, respectively, the air base is 682 m, and the camera focal length is 152.60 mm.

8-12. A pair of overlapping vertical photos is taken from a flying height of 981 m above ground with a 152.46-mm-focal-length camera. The *x* coordinates on the left photo of the base and top of a certain tree are 76.2 mm and 82.9 mm, respectively. On the right photo these *x′* coordinates are –13.8 mm and –16.1 mm, respectively. Determine the height of the tree.

8-13. A pair of overlapping vertical photos is taken from a flying height of 1273 m above the base of a radio tower. The *x* coordinates on the left photo of the top and base of the tower were 87.3 mm and 80.4 mm, respectively. On the right photo these *x′* coordinates were –11.6 mm and –10.2 mm, respectively. What is the approximate height of the tower?

8-14. The air base of a pair of overlapping vertical photos was determined to be 793 m. The focal length of the camera was 152.35 mm. The image coordinates of point *A*, whose elevation is 203 m above datum, were determined on the left photo as $x_a = 37.2$ mm and on the right photo as $x_a' = -52.9$ mm. What is the flying height above datum for the stereopair?

8-15. Repeat Prob. 8-14, except that the air base was 514 m, the camera focal length was 209.60 mm, and point *A*, whose elevation was 365 m above datum, had image coordinates of $x_a = 44.9$ mm on the left photo and $x_a' = -46.9$ mm on the right photo.

8-16. The images of two control points *A* and *B* appear in the overlap area of a pair of vertical photographs. The following photo coordinates and ground coordinates apply to points *A* and *B*. Calculate the air base of the stereopair, using Eq. (8-11).

	Left Photo Coordinates		Right Photo Coordinates		Ground Coordinates	
Point	*x*, mm	*y*, mm	*x′*, mm	*y′*, mm	*X*, m	*Y*, m
A	27.1	–96.7	–64.8	–96.7	78,164.6	27,996.4
B	–13.2	33.1	–86.8	33.0	78,250.2	27,351.6

8-17. Repeat Prob. 8-16, except that the photo coordinates and ground coordinates for points *A* and *B* were as follows:

	Left Photo Coordinates		Right Photo Coordinates		Ground Coordinates	
Point	x, mm	y, mm	x', mm	y', mm	X, m	Y, m
A	66.3	81.4	−33.5	81.5	102,055.8	35,781.1
B	39.6	−75.8	−53.3	−75.8	100,989.8	34,196.6

8-18. A pair of overlapping vertical photos was exposed with a camera having a 209.80-mm focal length. Calculate B and H from the following information on ground points D and E. [Hint: Set up Eq. (8-5) for point D and for point E, then solve simultaneously.]

Point	Elevation, m	Left Photo Coordinates	Right Photo Coordinates
D	577	$x_d = 23.6$ mm	$x'_d = -69.8$ mm
E	735	$x_e = 89.4$ mm	$x'_e = -8.3$ mm

8-19. Repeat Prob. 8-18, except that the camera focal length is 152.53 mm and the following information applies to points D and E.

Point	Elevation, m	Left Photo Coordinates	Right Photo Coordinates
D	546	$x_d = 71.8$ mm	$x'_d = -21.6$ mm
E	481	$x_e = 25.8$ mm	$x'_e = -63.9$ mm

8-20. A parallax wedge for use with a pocket stereoscope similar to that shown in Fig. 8-12 has height of graduations y equal to 160 mm. The lateral spacing between reference line and the graduated line is 45 mm at the top and 65 mm at the bottom. What is the vertical spacing of reference marks on the graduated line if the difference in parallax between adjacent graduations is 0.25 mm?

8-21. In Prob. 8-14, suppose that random errors were ±1 m in h and B, and ±0.05 mm in each of x_a and x_a'. What is the expected resultant error in the calculated value of H due to these random errors? (Assume the focal length to be error-free.)

8-22. In Prob. 8-15, suppose that random errors were ± 0.5 m in h_A, ± 2 m in B, and ±0.1 mm in both x_a and x_a'. What is the expected error in the calculated value of H due to these errors? (Assume the focal length to be error-free.)

8-23. In Prob. 8-12, assume that random errors existed in the amounts of ± 1 m in H and ±0.1 mm for each of the measured photo coordinates. What is the expected error in the calculated height of the tree due to these random errors?

CHAPTER 9

Elementary Methods of Planimetric Mapping for GIS

9-1 Introduction

This chapter describes elementary methods that can be used for compiling planimetric maps from vertical photographs and satellite images. These include (1) tracing with the use of reflection instruments, (2) *georeferencing* of digital imagery, (3) performing heads-up digitizing, (4) preparing *photomaps*, and (5) constructing *mosaics*.[1] Each of these techniques is relatively uncomplicated to perform and generally requires simpler and less expensive equipment compared to the rigorous photogrammetric mapping techniques that are presented in later chapters. These methods can have definite utility, depending upon the extent and required accuracy of planimetric mapping to be accomplished. For map revision over limited areas, and in many applications of geographic information systems (GISs), particularly those involving natural resource mapping, the high accuracy afforded by rigorous photogrammetric methods is often not necessary. By using appropriate elementary methods, substantial cost savings can be realized while obtaining planimetric information at acceptable accuracy.

The choice among the above-cited methods depends upon the purpose of the map or image product, extent of the area to be mapped, required accuracy, available imagery, and budget constraints. As an example, it may be necessary to include a recently constructed road or shopping center which is not shown on an otherwise satisfactory existing planimetric map. It would be expensive, and unnecessary, to prepare a new map of the area if these features could be satisfactorily superimposed onto the existing map. This type of planimetric map revision can readily be done using procedures described in this chapter.

The accuracies that can be achieved in planimetric mapping by using these methods are generally of a lower order than those attainable with stereoplotting instruments (see Chap. 12) or orthophoto processing (see Chap. 15). However, for some work, especially if ample care is exercised, suitable results can be achieved.

9-2 Planimetric Mapping With Reflection Instruments

A number of instruments have been devised that enable an operator to view an image superimposed with a map manuscript. The basic operation of these instruments is fundamentally the same. Photo control[2] points which have been plotted to scale on the map manuscript are used as a basis for orienting the image. By adjusting image scale, rotation, translation, and in some cases tilt or nonuniform image stretch, the images of the photo control points can be made to approximately coincide with the plotted control points on the manuscript. The degree of coincidence depends on a number of factors including the amount of terrain relief. Once the device is oriented, additional image features can be directly traced onto the manuscript by virtue of the operator's simultaneous view of the oriented image and the map.

The zoom transfer scope (ZTS) of Fig. 9-1 is a versatile reflection instrument which is used primarily for planimetric map revision. Although the use of ZTSs is declining, they can still be found being used at mapping agencies. With the ZTS, an aerial photograph or other hard-copy image is placed on a viewing stage and secured with clamps. Some ZTSs can accommodate imagery on either

transparent film or paper, and some can function in either monoscopic or stereoscopic mode. The map manuscript is placed on a horizontal working surface which can be viewed simultaneously with the image. The instrument can accommodate large differences in scale from photo to map by means of zoom optics which provide continuous variations in magnification from 1 to 7X. In addition, the anamorphic optical system of the ZTS enables different magnification ratios to be applied in the *x* and *y* directions. The ZTS also has the capability of removing the effects of nonorthogonal *affine* distortion (see Sec. C-6) in the image by means of rotating prisms. These features facilitate adjusting photo images to coincide with map points or control points. Once the image and map are adjusted, the manuscript is taped down and the operator then traces the desired features.

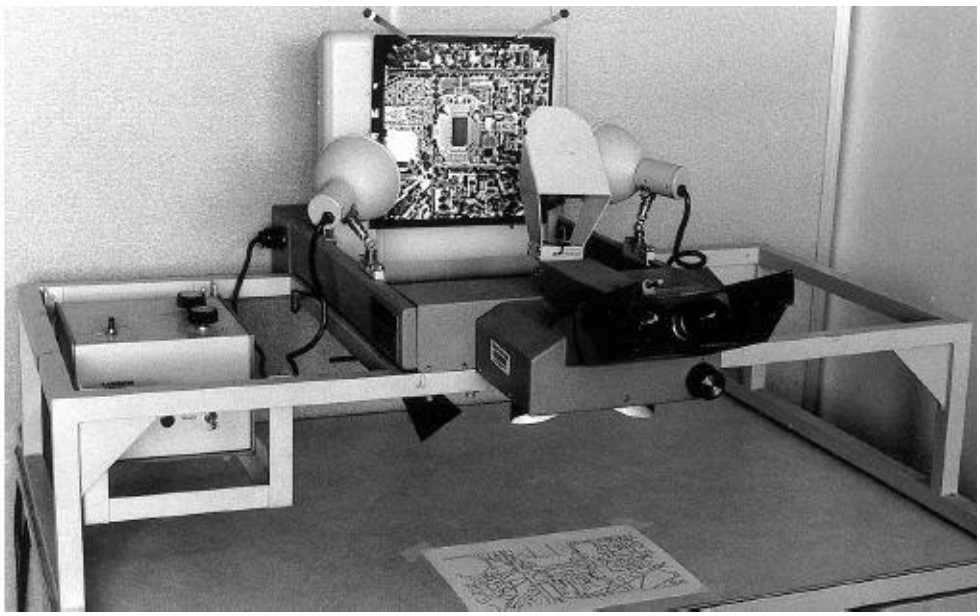


FIGURE 9-1 Zoom transfer scope. (*Courtesy University of Florida.*)

Of course, photographs are perspective projections, and as described in Chap. 6, they contain scale variations due to tilt and relief. Therefore, in compiling maps using reflection instruments, traced features will contain any scale variations that exist in the photo images. However, if relief is moderate, and relatively small areas are traced that lie within a perimeter of control points, the resulting map will be sufficiently accurate to satisfy many purposes.

9-3 Georeferencing of Digital Imagery

A digital image as described in Sec. 2-12 can be obtained through a number of sources including satellite imaging sensors, digital cameras, and scanned aerial photographs. In its rudimentary form, a digital image bears no relationship to a ground coordinate reference system. Rather, its coordinate basis consists of integer column and row numbers which specify a pixel's location within a rectangular image array. *Georeferencing*, sometimes called *ground registration*, is a technique whereby a digital image is processed so that the columns and rows of the resulting product are aligned with north and east in a ground coordinate system. Some authors, particularly those in the remote-sensing field, refer to georeferencing as *rectification;* however, in the context of photogrammetry, the term *rectification* is reserved for the process of removing the effects of tilt from an aerial photograph. (See Secs. 10-11 through 10-15.)

The process of georeferencing an image involves two fundamental steps: (1) computing the

parameters of a two-dimensional coordinate transformation (see App. C) which relates the digital image to the ground system and (2) filling an array, which is aligned with the ground coordinate system, with the appropriate digital numbers that quantify the brightness of the ground at the corresponding locations.

In the first step, a number of ground control points (GCPs) are selected which can be identified in the image and for which accurate ground coordinates are available. The column and row image coordinates of each GCP are obtained and subsequently related to the ground coordinates. The simplest method for obtaining image coordinates of a GCP is to display the image on a computer screen. Then by using a mouse, a cursor is guided to the location of the image point and recorded by clicking on it. Many image manipulation programs are capable of this rudimentary operation.

The coordinate transformation in this first step of the georeferencing process *converts from ground coordinates (x and y) to image coordinates (X and Y)*. At first, the direction of this conversion (from ground to image) may seem backward. However, for reasons which will become clear in this section, the transformation must be performed in this manner. Furthermore, the use of lowercase *x* and *y* for ground coordinates and uppercase *X* and *Y* for image coordinates seems to be a reversal from the usual convention. Actually, this is appropriate when one looks at the uppercase/lowercase convention from a different point of view. In the coordinate transformation equations presented in App. C, lowercase variables are used for coordinates in the "from" system (the initial system), and uppercase variables are used for coordinates in the "to" system (the final system). For georeferencing, since the ground system is the "from" system, lowercase *x* and *y* are used for its coordinates; and since the image system is the "to" system, uppercase *X* and *Y* are used for its coordinates. Image coordinates of the common points will serve as control for the transformation, and the resulting parameters will give the relationship from ground to image coordinates. Any two-dimensional coordinate transformation could be used, but the conformal and affine are most often employed because of their convenience and suitability. For illustration purposes here, the two-dimensional conformal coordinate transformation will be used. Equations (C-12) express this transformation relationship, and for convenience they are repeated here as Eq. (9-1). [Note that the variable names have been changed to relate to the coordinate system designations and specific conversion process (noted in italics) above.]

$$X = ax - by + T_X$$

$$Y = ay + bx + T_Y \tag{9-1}$$

In Eq. (9-1), *x* and *y* are coordinates of points in the ground system; *X* and *Y* are coordinates of points in the image that have been obtained by converting from their column and row values; and *a*, *b*, T_X, and T_Y are parameters which are determined during this first step. Once computed, these parameters are used to transform coordinates of additional points from the ground system to the image system. The reason for this arrangement of the transformation will become apparent in the description of the second step, which follows.

During the second step of georeferencing, an image that is aligned with the ground coordinate system is produced. To understand this step, it is useful to visualize the ground as being divided into a rectangular grid of individual elements called *groundels* (ground elements), analogous to pixels (picture elements) of a digital image. The difference is that while pixels have no specific relationship to the ground, groundels are arranged at a nominal spacing in a grid which is parallel to the ground coordinate system. For each groundel, the *x* and *y* coordinates of its center point are transformed, based on the parameters computed in the first step, into corresponding image coordinates. The

brightness value from the image at this corresponding location is then inserted into the groundel array. This involves the process of *resampling*, which is covered in App. E. After this process has been applied to each groundel, the georeferenced image is contained in the array of groundels.

To illustrate the above-described process of georeferencing, refer to Figs. 9-2*a* and *b*. Figure 9-2*a* represents an 8-pixel by 8-pixel digital image prior to georeferencing. Note that this image is represented as a square, nonrotated image, which is how it might appear if displayed on a computer screen. The solid-line grid of Fig. 9-2*b* represents an array of groundels which are nominally aligned with a ground coordinate system *xy*. Nominal ground coordinates are associated with the centers of the groundels, shown as small dots in Fig. 9-2*b*. The dashed lines represent the digital image as it would appear if properly aligned with the ground system. Four control points (1, 2, 3, and 4) having coordinates in the ground *xy* system appear in the image. Coordinates of these control points in both the image system of Fig. 9-2*a* and the ground system are listed in Table 9-1. Note that image coordinates, which correspond to pixel centers [indicated by crosses in Fig. 9-2*a*], are specified to the nearest whole row and column, which is typical of most georeferencing applications. More elaborate techniques are available which can yield accurate fractional row and column coordinates; however, these techniques are not discussed here.

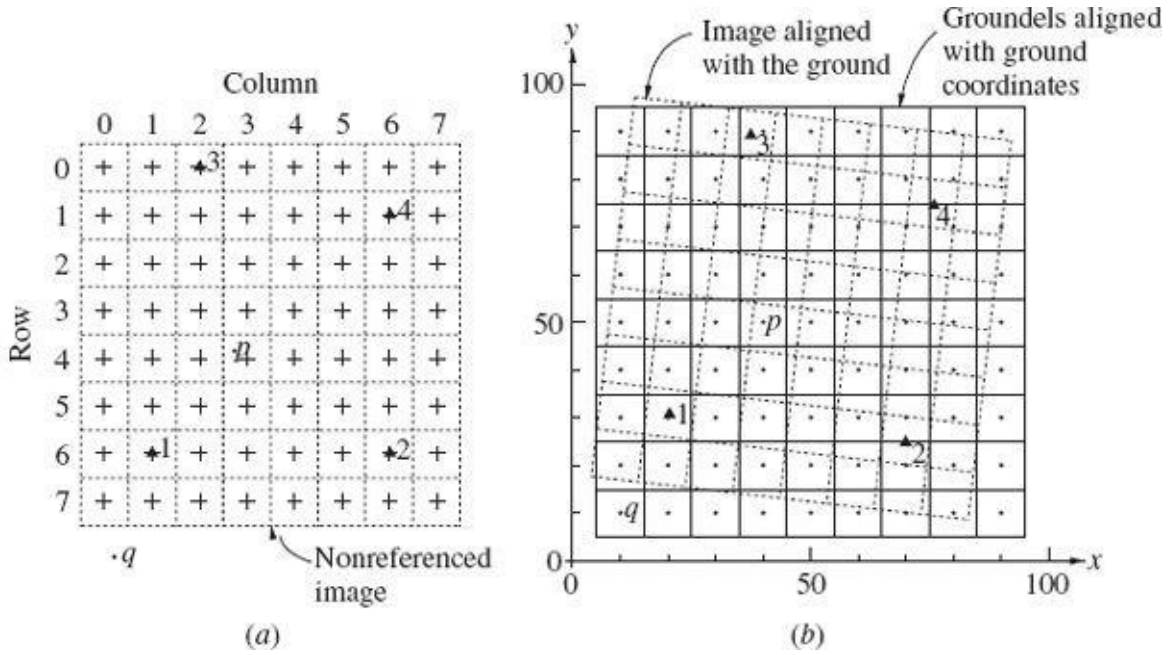


FIGURE 9-2 (*a*) Schematic representation of nonreferenced digital image. (*b*) Schematic representation of georeferenced image (solid lines) and orientation of image aligned with the ground (dashed lines).

	Ground Coordinates		Image Coordinates	
Point	*x*, m	*y*, m	Column	Row
1	20.4	30.6	1	6
2	70.1	24.9	6	6
3	37.1	89.3	2	0
4	75.8	74.4	6	1

TABLE 9-1 Ground and Image Coordinates for Control Points of Fig. 9-2

To prepare for the coordinate transformation which will yield the mathematical relationship between the image and ground coordinates, it is useful to convert the column and row coordinates of image pixels to the more conventional *XY* system. Note that in Fig. 9-2*a* the column coordinates increase from left to right, similar to conventional *X* coordinates; however the row coordinates increase from top to bottom, which is in the opposite direction from conventional *Y* coordinates. Because of this configuration, using the column coordinate as the abscissa and the row coordinate as the ordinate results in a "left-handed" coordinate system, as opposed to the more commonly used "right-handed" system, which is typical of ground coordinate systems. Transformation of a left-handed system to a right-handed system can cause mathematical inconsistencies in certain coordinate transformations, such as the conformal transformation. This inconsistency can be remedied by a number of approaches.

One approach is to use the row coordinate as the abscissa and the column coordinate as the ordinate. This results in a right-handed coordinate system which is rotated 90° clockwise. Another approach is to simply use a transformation that is able to account for the difference between left- and right-handed coordinate systems. The affine transformation has this capability by virtue of its separate *x* and *y* scale factors (see Sec. C-6), one of which will be positive and the other, negative. A third method is to convert the image coordinates from a left-handed system with the origin in the upper left corner to a right-handed system with its origin in the lower left corner. This can be accomplished by selecting the column coordinate as the abscissa and calculating ordinate values by subtracting row coordinates from the maximum row number.

Assume that the two-dimensional conformal transformation will be used to relate the ground coordinate system of Fig. 9-2*b* to the image coordinate system of Fig. 9-2*a*. Using image coordinates from Table 9-1, a right-handed *XY* coordinate system will be created using the third method just described. The maximum row number of the image is 7; therefore the original row coordinates will be subtracted from 7 to obtain *Y* coordinates. The resulting *XY* image coordinates are listed in Table 9-2.

	Original Image Coordinates		Converted Image Coordinates	
Point	Column	Row	X = Column	$Y = 7 - $ Row
1	1	6	1	1
2	6	6	6	1
3	2	0	2	7
4	6	1	6	6

TABLE 9-2 Column and Row Image Coordinates of Table 9-1 Converted to *XY* Coordinates

To transform from the *xy* ground coordinate system to the *XY* converted image coordinate system, the following equations [of the form of Eq. (9-1)] are written:

$$
\begin{aligned}
20.4a - 30.6b + T_X &= 1 + V_{X_1} \\
30.6a + 20.4b + T_Y &= 1 + V_{Y_1} \\
70.1a - 24.9b + T_X &= 6 + V_{X_2} \\
24.9a + 70.1b + T_Y &= 1 + V_{Y_2} \\
37.1a - 89.3b + T_X &= 2 + V_{X_3} \\
89.3a + 37.1b + T_Y &= 7 + V_{Y_3} \\
75.8a - 74.4b + T_X &= 6 + V_{X_4} \\
74.4a + 75.8b + T_Y &= 6 + V_{Y_4}
\end{aligned}
\tag{9-2}
$$

Using the least squares techniques presented in Secs. B-9 and C-5, the equations are solved for the most probable values of a, b, T_X, and T_Y, which are

$$
\begin{aligned}
a &= 0.09921 \qquad & T_X &= -0.67 \\
b &= 0.01148 \qquad & T_Y &= -2.27
\end{aligned}
$$

Once these transformation parameters have been computed, the georeferencing is completed by filling the groundel array with brightness values from the digital image. Referring to Fig. 9-2*b*, which contains a 9 × 9 groundel grid, this involves 81 separate applications of the above transformation, each followed by access of the brightness value from the image through resampling. For example, point *p* in Fig. 9-2*b* has ground coordinates of $x = 40$ and $y = 50$, and its image coordinates X and Y, by substitution into Eqs. (9-1), are

$$
\begin{aligned}
X &= 40a - 50b + T_X = 40(0.09921) - 50(0.01148) + (-0.67) = 2.7 \\
Y &= 50a + 40b + T_Y = 50(0.09921) + 40(0.01148) + (-2.27) = 3.1
\end{aligned}
$$

Expressing the location in column and row coordinates gives

$$
\begin{aligned}
\text{Column} &= X = 2.7 \\
\text{Row} &= 7 - Y = 3.9
\end{aligned}
$$

This column and row location is indicated by the position of point *p* in Fig. 9-2*a*. This position falls within the area of the pixel at column = 3 and row = 4, so the brightness associated with that particular image pixel could be copied into the corresponding groundel at $x = 40$ and $y = 50$. This is the *nearest-neighbor* method of resampling, which is described in App. E. Alternatively, *bilinear* or *bicubic* interpolation could be used for the resampling.

It is possible for some of the groundels to fall outside the limits of the digital image, as is the case with point *q*. From Fig. 9-2*b*, point *q* corresponds to the groundel at coordinates $x = 10$ and $y = 10$. Transforming these coordinates into the *XY* image system gives

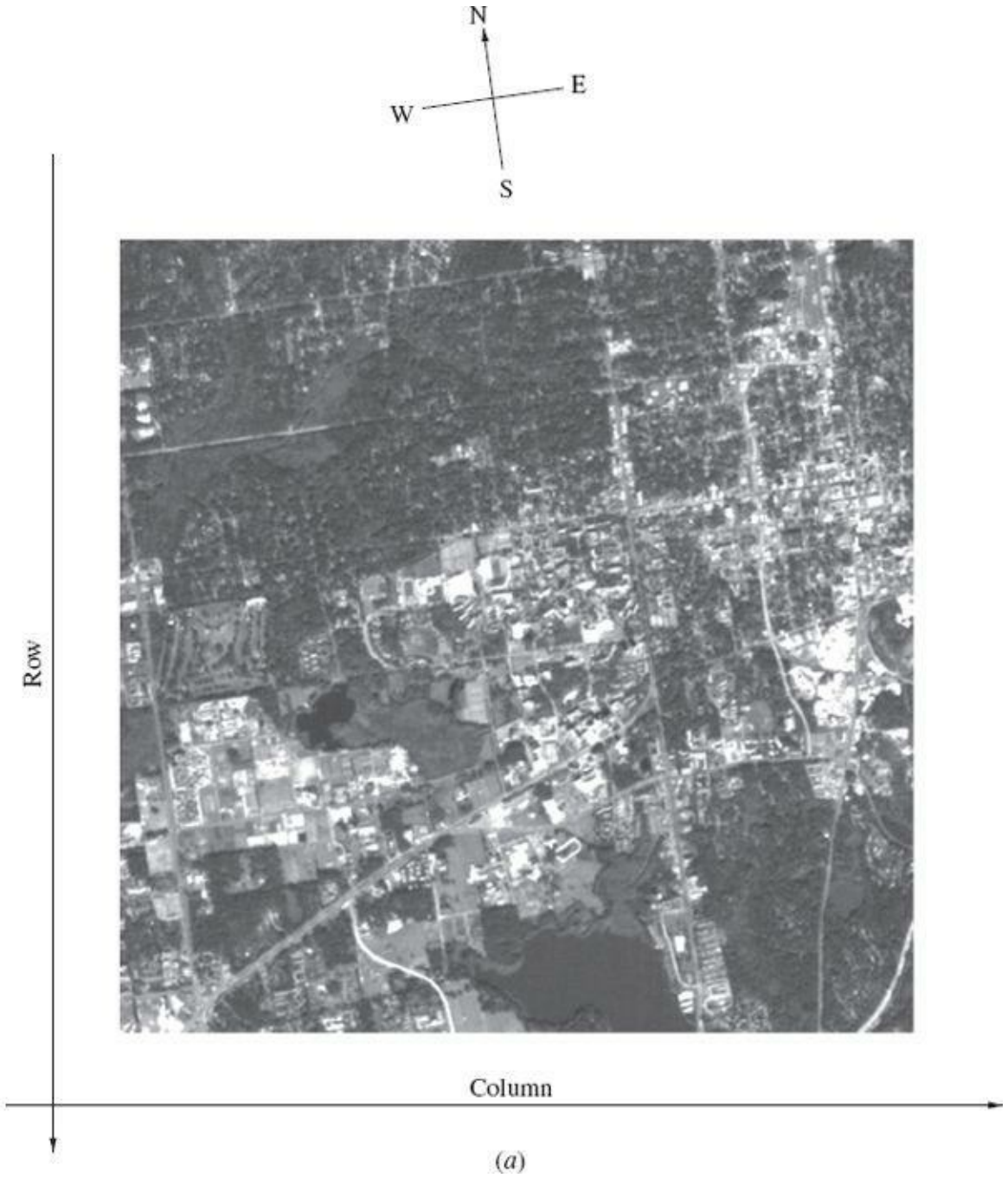
$$
\begin{aligned}
X &= 10a - 10b + T_X = 10(0.09921) - 10(0.01148) + (-0.67) = 0.2 \\
Y &= 10a + 10b + T_Y = 10(0.09921) + 10(0.01148) + (-2.27) = -1.2
\end{aligned}
$$

Expressing this location in column and row coordinates gives

$$\text{Column} = X = 0.2$$
$$\text{Row} = 7 - Y = 8.2$$

Notice that since the row coordinate of 8.2 is beyond the maximum row number of 7, there is no brightness value from the image corresponding to this particular groundel. In this case, it is appropriate to use a default value of 0 for the groundel brightness value.

Figure 9-3 shows the result of georeferencing as applied to a satellite image. Figure 9-3*a* shows the nonreferenced image. Notice that even though the true directions of streets in the area correspond to cardinal directions (north, south, east, and west), the general directions of streets appear to be rotated approximately 10° in the image. Figure 9-3*b* shows the resulting georeferenced image where the directions of streets now properly correspond to cardinal directions.



(*a*)

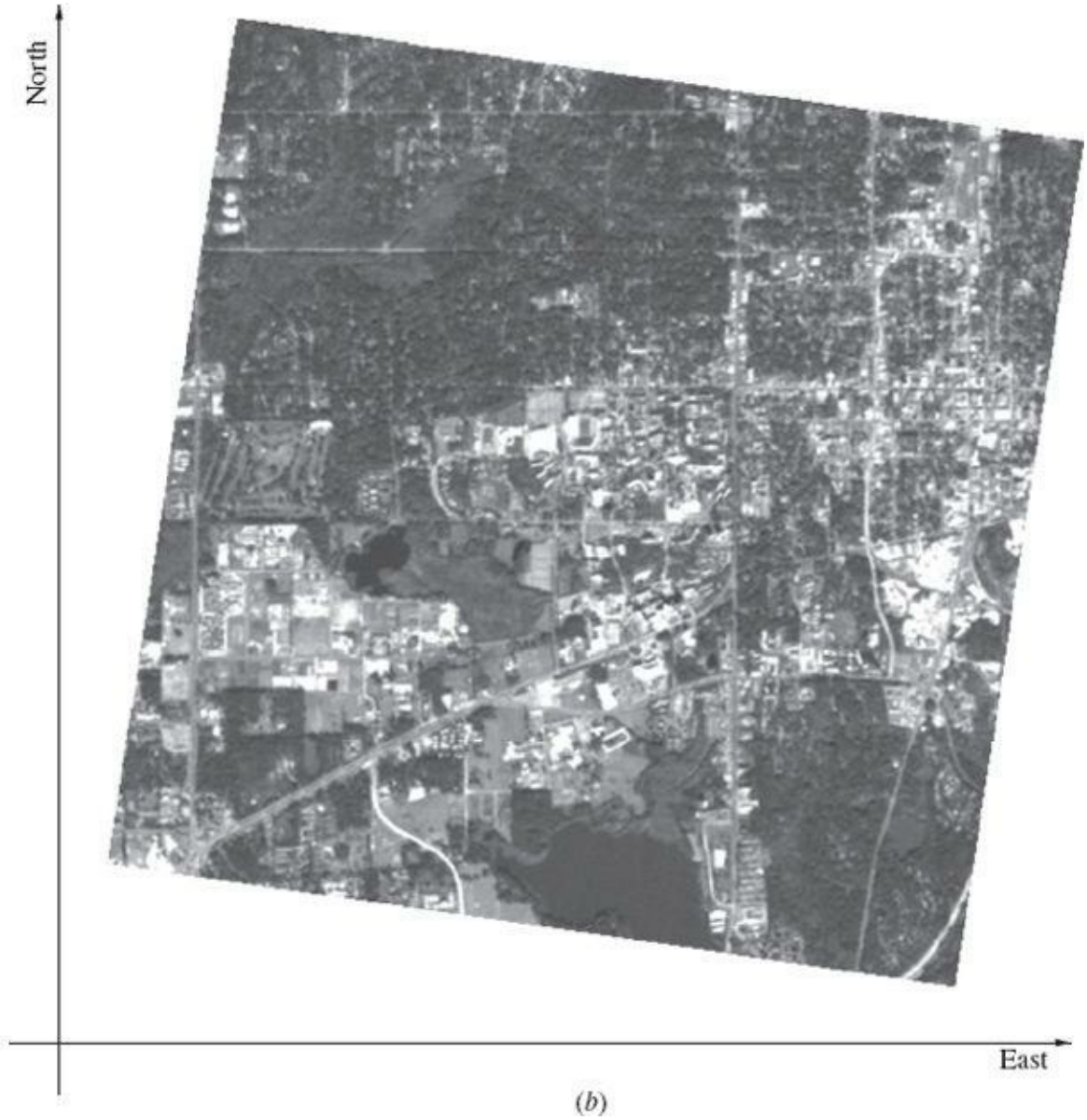


FIGURE 9-3 (*a*) Nonreferenced satellite image. (*b*) Georeferenced image. (*Courtesy University of Florida*.)

When georeferencing digital imagery, it is beneficial to choose groundel dimensions that are consistent with the pixel dimensions of the image. For example, when georeferencing a satellite image with 10-m pixels, the groundel size should also be 10 m. This is not an absolute necessity, but if the dimensions differ by more than a factor of about 1.5, more elaborate resampling methods may be necessary.

When choosing ground control points for georeferencing, it is important to select a well-distributed pattern of points, preferably with a point in each of the four corners as well as additional points spaced uniformly throughout the interior. It is particularly important to avoid having all control points in a small cluster in one area of the image. In addition, if the affine or projective transformation is used, it is critical that not all the control points lie along a straight line; otherwise, the transformation becomes ill-conditioned and yields poor results. Depending on the transformation used and the required geometric accuracy, the number of control points per image should be roughly 8 to 15. With a smaller number of control points, the accuracy will generally be lower and it will be more difficult to detect errors that may be present in the coordinates of one or more points. For the highest accuracy, a large number of control points are always desirable; however, beyond about 15 points, the

marginal benefit of additional points seldom outweighs the extra effort involved.

The geometric accuracy of a georeferenced image can be assessed through evaluation of the residuals from the transformation, which should always be computed and checked. Large residuals may indicate errors in measured image or ground coordinates. Many computer software packages calculate the *root mean square* (RMS) of the residuals, which gives a nominal indication of accuracy. This value can be misleading, however, since the probability of an actual point's position being within the two-dimensional RMS error region is only about 40 percent. To increase the level of confidence to about 95 percent, the RMS value should be multiplied by 2.5.

It is possible to compute a transformation with small residuals but still have low accuracy in terms of absolute position. This would occur if, for example, the ground coordinates were scaled from a map that contains a significant systematic error. Perhaps the map was based on the NAD27 datum, but the resulting georeferenced image is assumed to be in the NAD83 datum (see Sec. 5-7). To disclose systematic errors of this type, it is advisable to perform a field test on a randomly selected sample of image points (checkpoints) using an independent source of higher accuracy, such as a global positioning system (GPS) survey. The magnitudes of discrepancies between coordinates as obtained from the georeferenced image and those obtained by the higher-accuracy method can be used to assess accuracy.

9-4 Heads-Up Digitizing

Heads-up digitizing is a term used to describe the process where features, which are visible on a georeferenced digital image displayed on a computer screen, are digitized directly with a mouse. The human operator views the digital image simultaneously with the digitized features on the computer screen. Since the operator's gaze remains on the computer screen throughout, the operation is referred to as *heads-up*.

Heads-up digitizing requires software that is capable of displaying a georeferenced image in the proper coordinate reference frame and enables simultaneous on-screen digitizing. Many GIS software packages are capable of this operation as well as some CAD packages. A key advantage of heads-up digitizing is that the points and lines associated with the digitized features are displayed as an overlay on the digital image, and if mistakes are made while pointing on the features, they can be readily detected and subsequently remeasured.

9-5 Photomaps

Photomaps are simply aerial photos that are used directly as planimetric map substitutes. The photos were traditionally brought to some desired average scale by enlargement or reduction by projection printing (see Sec. 2-11), but are now commonly plotted on a large format printer from a digital file. Title information, place names, and other data may be superimposed on the photos in the same way that it is done on maps. Photomaps may be prepared from single aerial photos, or they may be made by piecing together two or more individual overlapping photos to form a single continuous composite picture. These composites are commonly referred to as *mosaics* and are described in Sec. 9-6.

Photomaps are similar to standard maps in many respects, but they have a number of definite advantages over maps. Photomaps show relative planimetric locations of a virtually unlimited number of objects, whereas features on maps—which are shown with lines and symbols—are limited by what was produced by the mapmaker. Photomaps of large areas can be prepared in much less time and at considerably lower cost than maps. Photomaps are easily understood and interpreted by people

without photogrammetry backgrounds because objects are shown by their images. For this reason they are very useful in describing proposed construction or existing conditions to members of the general public, who may be confused by the same representations on a map.

Photomaps have one serious disadvantage—they are not true planimetric representations. Rather, they are constructed from perspective photographs, which are subject to image displacements and scale variations. The most serious image displacements and scale variations are caused by variations in the terrain elevation and tilting of the camera axis. These displacements are generally most severe at the edges of the photographs. Of course, some small errors result from printer distortion of digital images (or shrinkage or expansion of the photo papers), camera lens imperfections, and atmospheric refraction, but these are generally negligible.

The effects of tilt can be eliminated by rectifying the photograph. (Rectification is described in Secs. 10-11 through 10-15.) Rectification does not remove the effects of topographic relief, however. Therefore the scale of the photomap is never constant throughout unless the terrain is perfectly level. Relief displacements can be minimized by increasing the flying height, while at the same time using a longer-focal-length camera to compensate for the decrease in scale. In measuring distances or directions from a photomap, it must be remembered that, due to image displacements, the scaled values will not be true. They are often used for qualitative studies only, and in that case slight planimetric inaccuracies caused by image displacements are of little consequence.

Because of their many advantages, photomaps are quite widely used. Their value is perhaps most appreciated in the field of planning, both in land-use planning and in planning for engineering projects. A photomap that shows an area completely and comprehensively can be rapidly and economically prepared. All critical features in the area that could affect the project can then be interpreted and taken into account. Alternative plans can be conveniently investigated, including considerations of soil types, drainage patterns, land-use and associated right-of-way costs, etc. As a result of this type of detailed study, the best overall plan is finally adopted.

Photomaps are valuable in numerous other miscellaneous areas besides planning. They are used to study geologic features, to inventory natural resources, to record the growth of cities and large institutions, to monitor construction activities at intervals of time, to record property boundaries, etc. They are also used as planimetric map substitutes for many engineering projects. Highway departments, for example, that are engaged in preparing plans for extensive construction projects frequently use photomaps to replace planimetric surveys. This not only eliminates much of the ground surveying but also does away with the office work associated with planimetric mapping. Design drawings and construction specifications are superimposed directly over the photomap. Used in that manner, these products have resulted in tremendous savings in time and cost, and they have yielded completely satisfactory results.

9-6 Mosaics

If a single photo does not contain extensive enough coverage to serve as a photomap of an area, an aerial mosaic may be prepared. Traditionally, mosaics were constructed manually from hard-copy paper prints, but in this day and age digital mosaics are more common. Whether accomplished manually or digitally, mosaics are constructed from a block of overlapping photographs which are trimmed and joined, much as cloth patches are stitched together to form a quilt. A special type of digital mosaic, known as a *composite orthophoto*, provides the most geometrically accurate image product available; however, its production is far more complex than that of a simple mosaic. Digital orthophotos are described in Sec. 13-8.

Aerial mosaics generally fall into three classes: *controlled*, *semicontrolled*, and *uncontrolled*. A

controlled mosaic is the most accurate of the three classes. In the manual process, this type of mosaic is prepared from photographs that have been rectified and ratioed; i.e., all prints are made into equivalent vertical photographs that have the same nominal scale. In assembling the mosaic, image positions of common features on adjacent photos are matched as closely as possible. To increase the overall accuracy of the assembly, a plot of control points is prepared at the same scale as the ratioed photos. Then in piecing the photos together to form the mosaic, the control point images are also matched to their corresponding plotted control points to constrain the positions of the photos. Controlled digital mosaics have similar characteristics, but they are prepared in an analytical process similar to georeferencing, as discussed in Sec. 9-3. Along the edges between adjacent photos, images of features are aligned to the extent possible, although they will seldom line up exactly. These residual misalignments exist primarily because of relief displacements.

An uncontrolled mosaic is prepared by simply matching the image details of adjacent photos. There is no ground control, and aerial photographs that have not been rectified or ratioed are used. Uncontrolled mosaics are more easily and quickly prepared than controlled mosaics. They are not as accurate as controlled mosaics, but for many qualitative uses they are completely satisfactory.

Semicontrolled mosaics are assembled by utilizing some combinations of the specifications for controlled and uncontrolled mosaics. A semicontrolled mosaic may be prepared, for example, by using ground control but employing photos that have not been rectified or ratioed. The other combination would be to use rectified and ratioed photos but no ground control. Semicontrolled mosaics are a compromise between economy and accuracy. The mosaic of Fig. 9-4 is a semi-controlled mosaic prepared from nonrectified photos but assembled to fit U.S. Geological Survey quadrangle maps.

FIGURE 9-4 Semicontrolled mosaic showing the entire San Francisco Bay area. This mosaic consists of more than 2000 individual aerial photographs. (*Copyright photo, courtesy Pacific Resources, Inc.*)

9-7 Uncontrolled Digital Mosaics

As noted above, a fundamental operation performed in the production of mosaics is the matching of images from adjacent photographs. In the manual process using hard-copy photographs, this is done through trial and error by shifting and rotating the photos slightly until optimal alignment is achieved. As previously mentioned, perfect alignment is seldom achieved due to relief displacements and tilt displacements of the images. With digital images, the matching procedure involves identifying *tie points* in the overlap areas between adjacent photos. To match a pair of photos, coordinates of the

common tie points are measured in both photos, and a two-dimensional coordinate transformation is performed to define the geometric relationship between the photographs. This process is illustrated in Fig. 9-5, which shows two adjacent photographs in superposition. In this figure, the overlap area between photos 1 and 2 contains four tie points, *a*, *b*, *c*, and *d*. In the separated photographs shown at the bottom of the figure, each of these tie points has row and column coordinates in the digital image coordinate system of each photo. For instance, point *a* has coordinates C_{1_a}, R_{1_a} in photo 1 and C_{2_a}, R_{2_a} in photo 2. By selecting the coordinates of the tie points from photo 2 as control, a two-dimensional coordinate transformation can be computed that determines the parameters for transforming photo 1 coordinates to photo 2 coordinates. In the example of Fig. 9-5, since the tie points lie nearly in a straight line, a conformal transformation should be used because more complex transformations, such as affine or projective, become mathematically unstable when the control configuration is linear or nearly so.

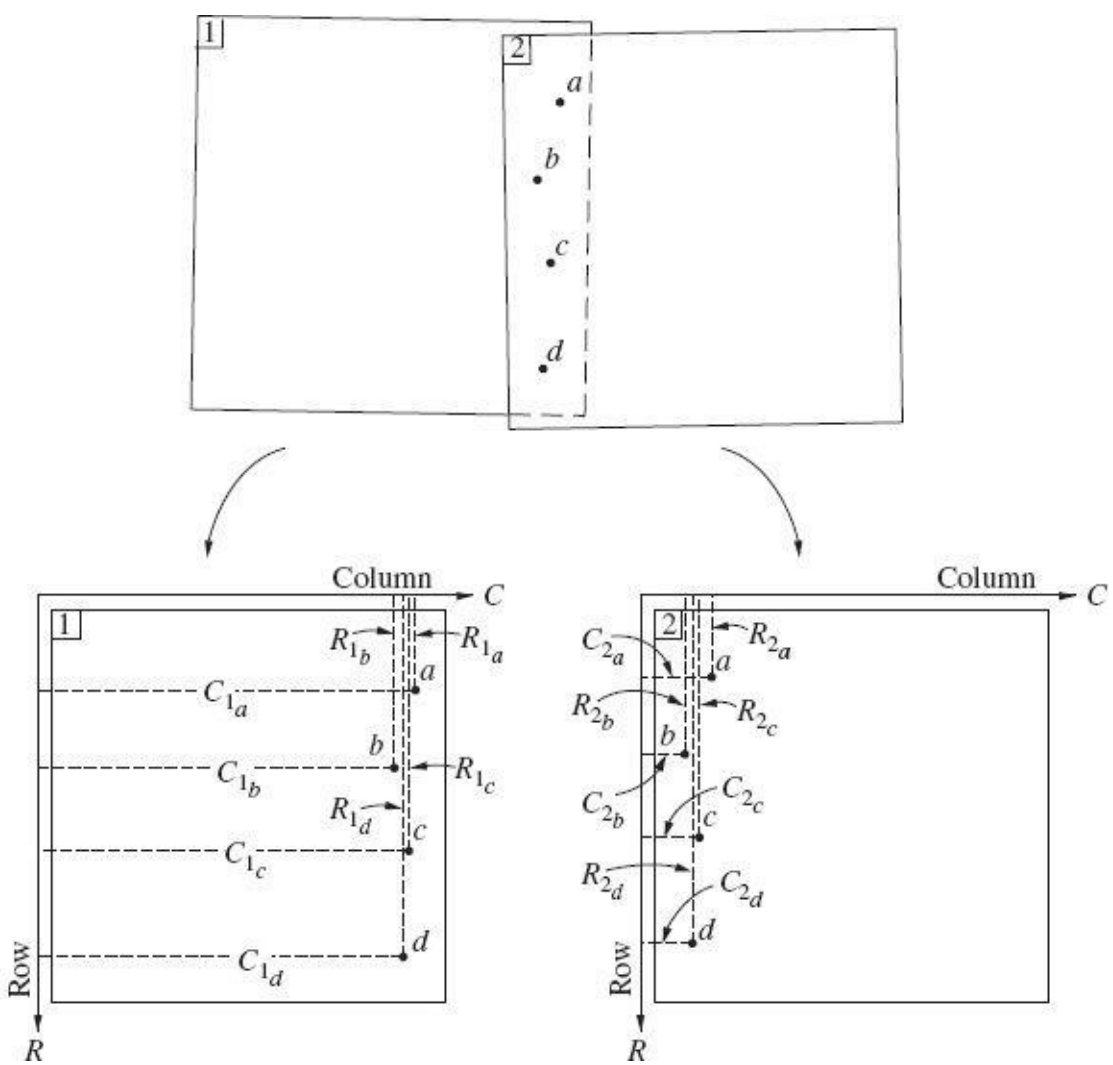


FIGURE 9-5 Use of tie points to join adjacent digital photographs.

After the transformation, the digital coverage of photo 1 can be extended to include the additional area covered by photo 2. This is done by increasing the number of columns in the photo 1 array to create sufficient storage space to accommodate additional pixels from the photo 2 array. The brightness values associated with these additional columns are initially undefined. Then, in a process similar to that of georeferencing, the coordinates of each pixel in the extended area of photo 1 are transformed to the coordinate system of photo 2, using the previously determined parameters.

Brightness values are then resampled from photo 2 and stored in the extended portion of the photo 1 array. The result is a single image which incorporates the coverage of both photos.

When two photographs are attached by the foregoing method, a choice can be made as to which one is used to provide the brightness values in the overlap area. The simplest approach is to retain the values from the photo 1 array. This is generally not the best approach for two reasons. First, the extreme right edge of photo 1 typically contains the image of a side fiducial mark, which obscures the underlying terrain. Second, since relief displacements are more severe at greater distances from the principal point of a vertical photo, images of features along the seam between photos are more likely to be disjointed. To minimize these problems, the common area should be divided approximately into two halves, with one-half of the brightness values coming from each photograph. If a natural edge such as a street or fence line runs through the overlap area from top to bottom, it can be used to define the seam between the two photos. This will generally result in a less noticeable seam in the resulting image.

Large, uncontrolled digital mosaics can be constructed by repeating this attachment process for each photograph in a strip or block. As more digital photos are added, the area of the mosaic grows until the entire region is covered. Successive attachments of photos create an accumulation of errors and can cause the digital mosaic to become significantly warped. This effect can be reduced somewhat by starting in the middle of the block and working outward in all directions.

9-8 Semicontrolled Digital Mosaics

As mentioned in Sec. 9-6, semicontrolled mosaics employ either rectified photos which can be joined (without control points) in the manner described in Sec. 9-7, or control points are used to constrain nonrectified photographs. The more likely situation associated with generation of a semicontrolled digital mosaic is to use nonrectified photographs in conjunction with a distributed configuration of control points. This situation is illustrated in Fig. 9-6. The figure shows 12 photographs (I through XII) arranged in a block formation with tie points (1 through 33) in the overlap areas between adjacent photos. Also, seven control points (*A* through *G*) are distributed about the block. Note that there is no single photograph with more than one control point; therefore, individual rectification of photographs is not possible. It is possible, however, to simultaneously transform points from all 12 photographs in a single least squares adjustment.

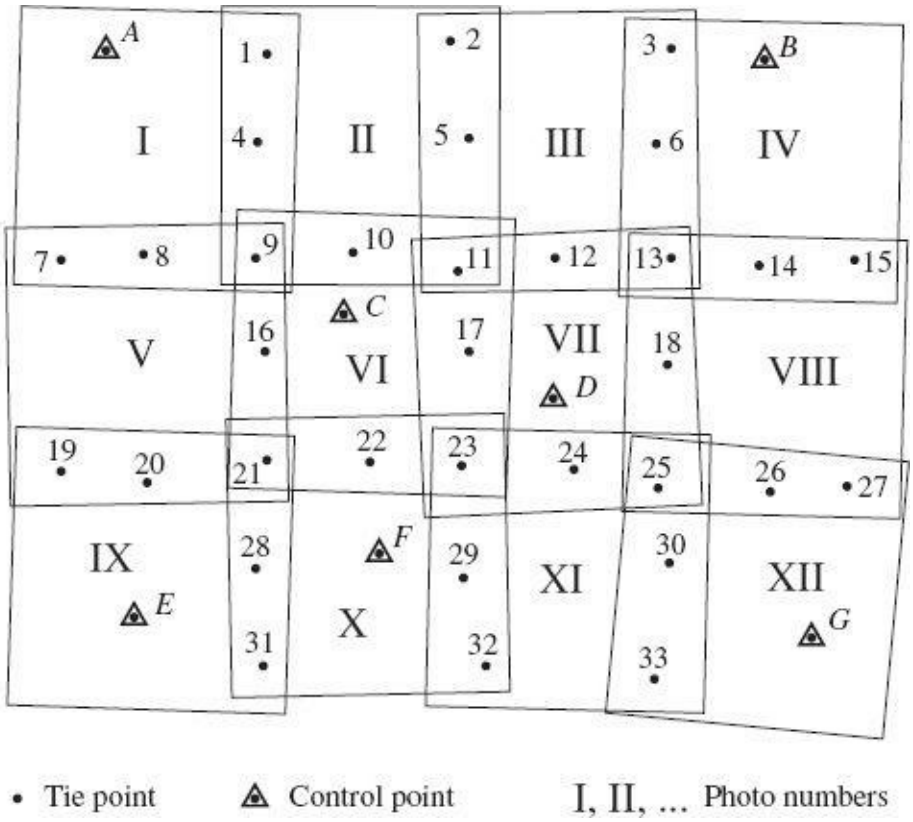


FIGURE 9-6 Control and tie point configuration for a semicontrolled digital mosaic.

Suppose that the two-dimensional conformal transformation is used to perform the adjustment. Equations (9-1) express the conformal transformation relationship between coordinates x and y in an arbitrary system and coordinates X and Y in the control system. In Eqs. (9-1), a, b, T_X, and T_Y are the four parameters of the transformation.

Since the block shown in Fig. 9-6 contains 12 individual photographs, 12 sets of four parameters (total of 48) are required to transform the individual photographs to the ground coordinate system. Calculation of the 48 unknown parameters requires a least squares adjustment involving two forms of observation equation pairs. The first type of observation equation pair is used for control points, and it has the same form as Eqs. (9-1) except that appropriate subscripts are included for the parameters and coordinates. As an example, note that in Fig. 9-6 control point C appears in photo VI. Digital photo coordinates x and y of the image of C in photo VI, which can be obtained from the row and column as described in Sec. 9-3, appear on the right-hand side of Eqs. (9-1); and X and Y ground coordinates of point C appear on the left. Also included in the right-hand side of Eqs. (9-1) are the parameters a, b, T_X, and T_Y which correspond to photo VI. Equations (9-3) express the observation equation pair which corresponds to the coordinates of control point C on photo VI, with appropriate subscripts.

$$X_C = a_{\text{VI}} x_{C_{\text{VI}}} - b_{\text{VI}} y_{C_{\text{VI}}} + T_{X_{\text{VI}}}$$

$$Y_C = a_{\text{VI}} y_{C_{\text{VI}}} + b_{\text{VI}} x_{C_{\text{VI}}} + T_{Y_{\text{VI}}} \tag{9-3}$$

Equations of this form are written for each control point and for each photo in which it appears. The second type of observation pair is for tie points. Since tie points have no ground coordinates associated with them, they cannot be used directly in Eqs. (9-1). Instead, the condition that the differences in ground coordinates X and Y for the common tie point based on the photo coordinates x

and y in two adjacent photographs must equal zero, is enforced. An example of this condition is given by Eqs. (9-4), which express the relationship for tie point 1 between photos I and II.

$$a_{\mathrm{I}}x_{1_{\mathrm{I}}} - b_{\mathrm{I}}y_{1_{\mathrm{I}}} + T_{X_{\mathrm{I}}} - (a_{\mathrm{II}}x_{1_{\mathrm{II}}} - b_{\mathrm{II}}y_{1_{\mathrm{II}}} + T_{X_{\mathrm{II}}}) = 0$$

$$a_{\mathrm{I}}y_{1_{\mathrm{I}}} + b_{\mathrm{I}}x_{1_{\mathrm{I}}} + T_{Y_{\mathrm{I}}} - (a_{\mathrm{II}}y_{1_{\mathrm{II}}} + b_{\mathrm{II}}x_{1_{\mathrm{II}}} + T_{Y_{\mathrm{II}}}) = 0 \tag{9-4}$$

A pair of equations of the form of Eqs. (9-4) is written for each occurrence of a tie point between a pair of photographs. Table 9-3 lists all the unique occurrences of tie points between adjacent photographs corresponding to Fig. 9-6. In this table, there are a total of 63 unique tie point connections between photographs, each of which contributes two equations of the form of Eqs. (9-4). Notice that tie points, such as point 1, which exist only on two photographs contribute only one pair of equations, whereas tie points, such as point 9, which exist on four photographs contribute six pairs of equations to the solution.

Point	Photo *i*	Photo *j*	Point	Photo *i*	Photo *j*	Point	Photo *i*	Photo *j*
1	I	II	13	III	VII	13	VII	VIII
4	I	II	13	III	VIII	18	VII	VIII
9	I	II	13	IV	VII	25	VII	VIII
7	I	V	13	IV	VIII	23	VII	X
8	I	V	14	IV	VIII	23	VII	XI
9	I	V	15	IV	VIII	24	VII	XI
9	I	VI	9	V	VI	25	VII	XI
2	II	III	16	V	VI	25	VII	XII
5	II	III	21	V	VI	25	VIII	XI
11	II	III	19	V	IX	25	VIII	XII
9	II	V	20	V	IX	26	VIII	XII
9	II	VI	21	V	IX	27	VIII	XII
10	II	VI	21	V	X	21	IX	X
11	II	VI	11	VI	VII	28	IX	X
11	II	VII	17	VI	VII	31	IX	X
3	III	IV	23	VI	VII	23	X	XI
6	III	IV	21	VI	IX	29	X	XI
13	III	IV	21	VI	X	32	X	XI
11	III	VI	22	VI	X	25	XI	XII
11	III	VII	23	VI	X	30	XI	XII
12	III	VII	23	VI	XI	33	XI	XII

TABLE 9-3 Connections Between Photos of Fig. 9-6 Based on Measured Tie Point Coordinates

After all the equations have been formed, the unknown parameters (a, b, T_X, and T_Y) associated with each photo are computed by least squares. These parameters specify the geometric relationships

between each photo and the ground system; however, the direction is the opposite of what is required. As described in Sec. 9-3 regarding georeferencing, coordinates of groundels must be transformed to the image coordinate system so that the brightness of the particular ground element can be determined. Since parameters a, b, T_X, and T_Y in Eqs. (9-2) transform from image coordinates to ground, the equation must be rearranged. To do this, equations can be written in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \end{bmatrix} \tag{9-5}$$

Solving Eq. (9-5) for the values of x and y results in Eq. (9-6).

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}^{-1} \begin{bmatrix} X - T_X \\ Y - T_Y \end{bmatrix} = \frac{1}{a^2 + b^2} \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} X - T_X \\ Y - T_Y \end{bmatrix} \tag{9-6}$$

By using the X and Y coordinates for each of the groundels in the area covered by the mosaic and the parameters of the corresponding photograph, Eq. (9-6) can be solved to obtain the position on the photo from which to extract the brightness values. The values are then resampled from the digital photo and placed in the groundel array. In areas of overlap between photographs, brightness values can be extracted from either photo. The choice can be made by the approach described in Sec. 9-7.

9-9 Controlled Digital Mosaics

As mentioned in Sec. 9-6, controlled mosaics are made from rectified and ratioed photographs in conjunction with ground control. With digital mosaics, this process begins with photo rectification, as described in Sec. 10-11. In this process control points are identified in each photograph, and the two-dimensional projective transformation is used to georeference the image, using the process described in Sec. 9-3. A minimum of four control points are needed, although more than four points are recommended in order to have redundancy and enable a least squares solution. Since each rectified digital photograph is georeferenced, the individual groundels will be directly related to ground coordinates. All that remains to create the controlled digital mosaic is to assemble the georeferenced images in a single groundel array that encompasses the full ground area of the mosaic. No further resampling or other significant processing is required, although choices must be made as to which brightness values to choose in overlapping areas.

An alternative approach can be used where there is insufficient ground control to rectify each digital photograph individually. It is a slight variation of the simultaneous coordinate transformation approach described in Sec. 9-8 which pertaines to semicontrolled digital mosaics. The difference is that instead of using a two-dimensional conformal or affine transformation, the projective transformation is used. This involves solving for all the individual sets of eight projective transformation parameters for the entire block simultaneously. Since the transformation is projective, the photographs will be rectified and ratioed as well as adjusted to ground control in a single operation.

A simultaneous projective transformation is more complicated than either the conformal or the affine because the equations are nonlinear. The approach of simplifying the equations into a linear form as discussed in Sec. C-9 [see Eqs. (C-55)] cannot be used for tie points since the X and Y coordinates appear on both sides of the equation. Performing a solution of the nonlinear form of the

projective equations requires a Taylor series expansion of Eqs. (C-54) as well as initial approximations for the unknown parameters. Equations (C-54) are repeated here as Eqs. (9-7) for convenience.

$$x = \frac{a_1 x + b_1 y + c_1}{a_3 x + b_3 y + 1}$$

$$y = \frac{a_2 x + b_2 y + c_2}{a_3 x + b_3 y + 1} \tag{9-7}$$

For near-vertical photographs, a simultaneous affine transformation can be used to obtain initial approximations, because the parameters a_3 and b_3 of Eqs. (9-7) are approximately zero. This causes the denominators in these equations to reduce to unity, and thus the projective Eqs. (9-7) are equivalent to the affine equations [see Eqs. (C-22)]. When the initial approximations have been determined, the simultaneous projective transformation equations (which have been linearized by Taylor's series) can then be formed, and iterated until the solution converges. This results in a solution for the projective parameters which specifies the geometric relationship between each photograph and the ground coordinate system.

As was the case with the simultaneous conformal transformation method described in Sec. 9-8, the computed parameters describe the relationship in the reverse sense from that required. To reverse Eqs. (9-7) so that the x and y photo coordinates are isolated on the left requires substantial algebraic manipulation. The algebraic steps required to reverse these equations are left to the reader. Equations (9-8) express the necessary relationship in reverse.

$$x = \frac{(b_2 - b_3 c_2)X + (b_3 c_1 - b_1)Y + b_1 c_2 - b_2 c_1}{(a_2 b_3 - a_3 b_2)X + (a_3 b_1 - a_1 b_3)Y + a_1 b_2 - a_2 b_1}$$

$$y = \frac{(a_3 c_2 - a_2)X + (a_1 - a_3 c_1)Y + a_2 c_1 - a_1 c_2}{(a_2 b_3 - a_3 b_2)X + (a_3 b_1 - a_1 b_3)Y + a_1 b_2 - a_2 b_1} \tag{9-8}$$

Some final comments must be made regarding the simultaneous projective transformation solution. The least squares solution is often sensitive to large amounts of relief in the project area. Since the solution involves tie points for which there is no known height, corrections cannot be made for their relief displacements. If heights in object space of the selected tie points vary substantially from mean terrain, the projective equations can become ill-conditioned, which can cause the least squares solution to fail. In this case, the only recourse is to rectify each digital photograph individually and assemble them as mentioned at the beginning of this section. In addition to the difficulties caused by substantial relief, simultaneous projective transformations require more control points throughout the photo block as well as more tie points along the edges than either the simultaneous conformal or affine transformation.

References

American Society for Photogrammetry and Remote Sensing: Manual of Photogrammetry, 5th ed., Bethesda, MD, 2004.

Derenyi, E. E.: "The Digital Transferscope," *Photogrammetric Engineering and Remote Sensing*, vol.

62, no. 6, 1996, p. 733.
Gao, J., and S. M. O'Leary: "The Role of Spatial Resolution in Quantifying SSC from Airborne Remotely Sensed Data," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 3, 1997, p. 267.
Homer, C. G., R. D. Ramsey, T. C. Edwards, Jr., and A. Falconer: "Landscape Cover-Type Modeling Using a Multi-Scene Thematic Mapper Mosaic," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 1, 1997, p. 59.
Jessiman, E. G., and M. R. Walsh: "An Approach to the Enhancement of Photomaps," *Canadian Surveyor*, vol. 30, no. 1, 1976, p. 11.
Kardoulas, N. G., A. C. Bird, and A. I. Lawan: "Geometric Correction of SPOT and Landsat Imagery: A Comparison of Map- and GPS-Derived Control Points," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 10, 1996, p. 1173.
Marsik, Z.: "Use of Rectified Photographs and Differentially Rectified Photographs for Photo Maps," *Canadian Surveyor*, vol. 25, 1971, p. 567.
Novak, K.: "Rectification of Digital Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 3, 1992, p. 339.
Steinwand, D. R., J. A. Hutchinson, and J. P. Snyder: "Map Projections for Global and Continental Data Sets and an Analysis of Pixel Distortion Caused by Reprojection," *Photogrammetric Engineering and Remote Sensing*, vol. 61, no. 12, 1995, p. 1487.
Vickers, E. W.: "Production Procedures for an Oversize Satellite Image Map," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 2, 1993, p. 247.

Problems

9-1. Describe the various techniques available for elementary planimetric mapping using vertical photographs.

9-2. Describe the zoom transfer scope and its advantages in planimetric mapping by direct tracing.

9-3. Discuss ground control requirements necessary for georeferencing.

9-4. Briefly discuss the advantage of heads-up digitizing over mapping with a tablet digitizer.

9-5. Discuss the advantages and disadvantages of photomaps and aerial mosaics as compared to conventional line and symbol maps.

9-6. Outline some of the uses of photomaps.

9-7. Define the three main classes of mosaics.

9-8. A photomap is to be prepared by enlarging a 23-cm-square-format aerial photo to a size of an 80-cm square. If a Lear jet was used to obtain the photo and a 152-mm-focal-length camera was carried to an altitude of 12.2 km above ground for the exposure, what will be the resulting scale of the photomap, and how many hectares will it cover?

9-9. A mosaic to serve as a map substitute for locating utilities will be prepared from digital photographs scanned at a 25-micrometer (25-μm) pixel size. The photographs will be taken with a

312-mm-focal-length camera having a 23-cm-square format. If 75-cm-diameter manholes must span 10 pixels on the digital mosaic, what must be the flying height above average ground for the photography?

9-10. Repeat Prob. 9-9, except the camera lens focal length is 152 mm, the photos will be scanned at a 12-μm pixel size, and 10-cm-wide highway stripes must span 2 pixels on the digital mosaic.

---

[1] In this context, compiling planimetric maps denotes the drawing of scaled diagrams which show planimetric features by means of lines and symbols or the production of images which portray planimetric features in picture form. Planimetric maps portray only horizontal position and give no information concerning elevations. The maps are prepared to some designated scale; thus all features are presumed to be shown in their true relative positions. Pure digital products such as computer-aided drafting (CAD) drawings and georeferenced images do not have a specific scale per se, but rather include explicit information regarding ground coordinates of features or pixels.

[2] Photo control, as described in detail in Chap. 16, consists of any discrete objects whose images appear on the photo and whose ground coordinates are known. There are several methods for determining the ground coordinates, including scaling from an existing map, direct field survey, or photogrammetric techniques such as aerotriangulation (described in Chap. 17).

CHAPTER 10
Tilted and Oblique Photographs

10-1 Introduction

In spite of level vials, inertial measurement devices, gimbals, and other stabilizing equipment, in practice it is impossible to maintain the optical axis of an airborne camera truly vertical. Unavoidable aircraft tilts cause photographs to be exposed with the camera axis tilted slightly from vertical, and the resulting pictures are called *tilted photographs*. If vertical photography is intended, the amount by which the optical axis deviates from vertical is usually less than 1° and rarely exceeds 3°. In some cases, aerial photography is purposely angled away from vertical. These types of images are classified as *high oblique* if the photograph contains the horizon, and *low oblique* otherwise.

Terrestrial photos are almost always taken from an oblique pose. *Horizontal* terrestrial photos are obtained if the camera axis is horizontal when the exposure is made. If a camera is available that has leveling capabilities, horizontal terrestrial photos can be obtained. Although, it is generally not practical to take horizontal photos because it is usually necessary to incline the camera axis up or down somewhat in order to center the object of interest in the field of view. Furthermore, no significant benefits result from taking horizontal photos.

Six independent parameters called the *elements of exterior orientation* (sometimes called EOPs) express the *spatial position* and *angular orientation* of a photograph. The spatial position is normally given by X_L, Y_L, and Z_L, the three-dimensional coordinates of the exposure station in a ground coordinate system. Commonly, Z_L is called H, the height above datum. Angular orientation is the amount and direction of tilt in the photo. Three angles are sufficient to define angular orientation, and in this book two different systems are described: (1) the *tilt-swing-azimuth* (t-s-α) system and (2) the *omega-phi-kappa* (ω-φ-κ) system. The omega-phi-kappa system possesses certain computational advantages over the tilt-swing-azimuth system and is therefore more widely used. The tilt-swing-azimuth system, however, is more easily understood and shall therefore be considered first.

10-2 Point Perspective

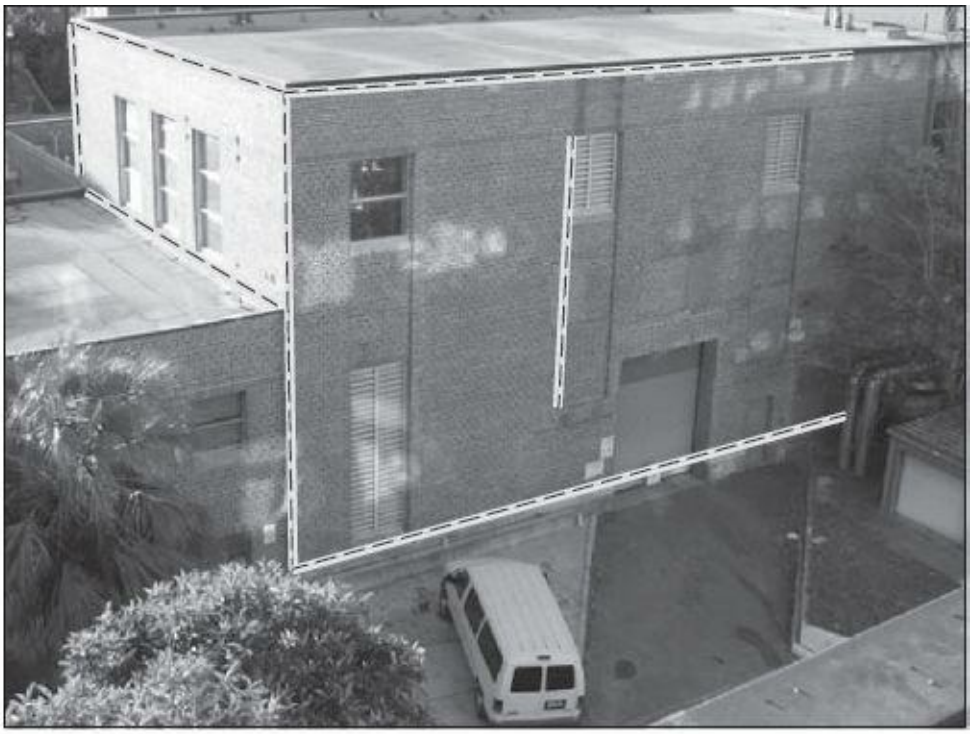
Perspective can be defined as how something is visually perceived under varying circumstances. In the context of frame imaging systems, *perspective projection* governs how 3D objects (the scene photographed) appear when projected onto a 2D surface (film or CCD array), making it the foundation for many of the principles of photogrammetry. The laws of perspective geometry can be used to derive information such as object scale and shape from tilted and oblique photographs and in turn derive information about the position and orientation of the photograph using knowledge of the objects being imaged. The collinearity equations mathematically describe the perspective projection from 3D object space to 2D image space (see App. D).

Some important properties of a perspective projection are that straight lines are preserved and projected lines that are parallel in object space intersect when projected onto the image plane (in fact, all lines intersect in projective space). More specifically, horizontal parallel lines intersect at the horizon, and vertical lines intersect at either the zenith or nadir in an image. The locations at which these lines intersect on images are called *vanishing points*. Figures 10-1, 10-2, and 10-3 illustrate one-, two-, and three-point perspective images of a building, respectively. Notice that in Fig. 10-1, the

horizontal lines at the left appear to converge toward a vanishing point in the distance, while the vertical lines appear parallel and the horizontal lines to the right appear parallel. For this figure, the optical axis of the camera was oriented perpendicular to the side of the building which resulted in one point perspective. In Fig. 10-2, there are two sets of horizontal parallel lines shown along two sides of the building. Both sets of lines, if extended, intersect at two points on the horizon to the left and right of the photo. The vertical lines shown in the photo appear parallel because the optical axis of the camera was horizontal (perpendicular to the vertical). It was not perpendicular to either side of the building however, which resulted in two point perspective. Finally, Fig. 10-3 shows three point perspective of the building from two different poses. In Fig. 10-3*a*, the horizontal lines also converge toward two points on the horizon, and since the camera is inclined down from horizontal, the vertical lines converge toward the nadir. Conversely, in Fig. 10-3*b*, since the camera is inclined up from horizontal, the vertical lines converge toward the zenith.

FIGURE 10-1 One-point perspective image.

FIGURE 10-2 Two-point perspective image.



(*a*)

(*b*)

FIGURE 10-3 (*a*) Three-point perspective image where vertical lines intersect at nadir. (*b*) Three-point perspective image where vertical lines intersect at zenith.

10-3 Angular Orientation in Tilt, Swing, and Azimuth

In Fig. 10-4, a tilted aerial photograph is depicted showing the tilt-swing-azimuth angular orientation parameters. In the figure, *L* is the exposure station and *o* is the principal point of the photo positive. Line *Ln* is a vertical line, *n* being the *photographic nadir point*, which occurs where the vertical line intersects the plane of the photograph. The extension of *Ln* intersects the ground surface at N_g, the *ground nadir point*, and it intersects the datum surface at N_d, the *datum nadir point*. Line *Lo* is the camera optical axis; its extension intersects the ground at P_g, the ground principal point, and it intersects the datum plane at P_d, the *datum principal point*. One of the orientation angles, *tilt*, is the angle *t* or *nLo* between the optical axis and the vertical line *Ln*. The tilt angle gives the magnitude of tilt of a photo.

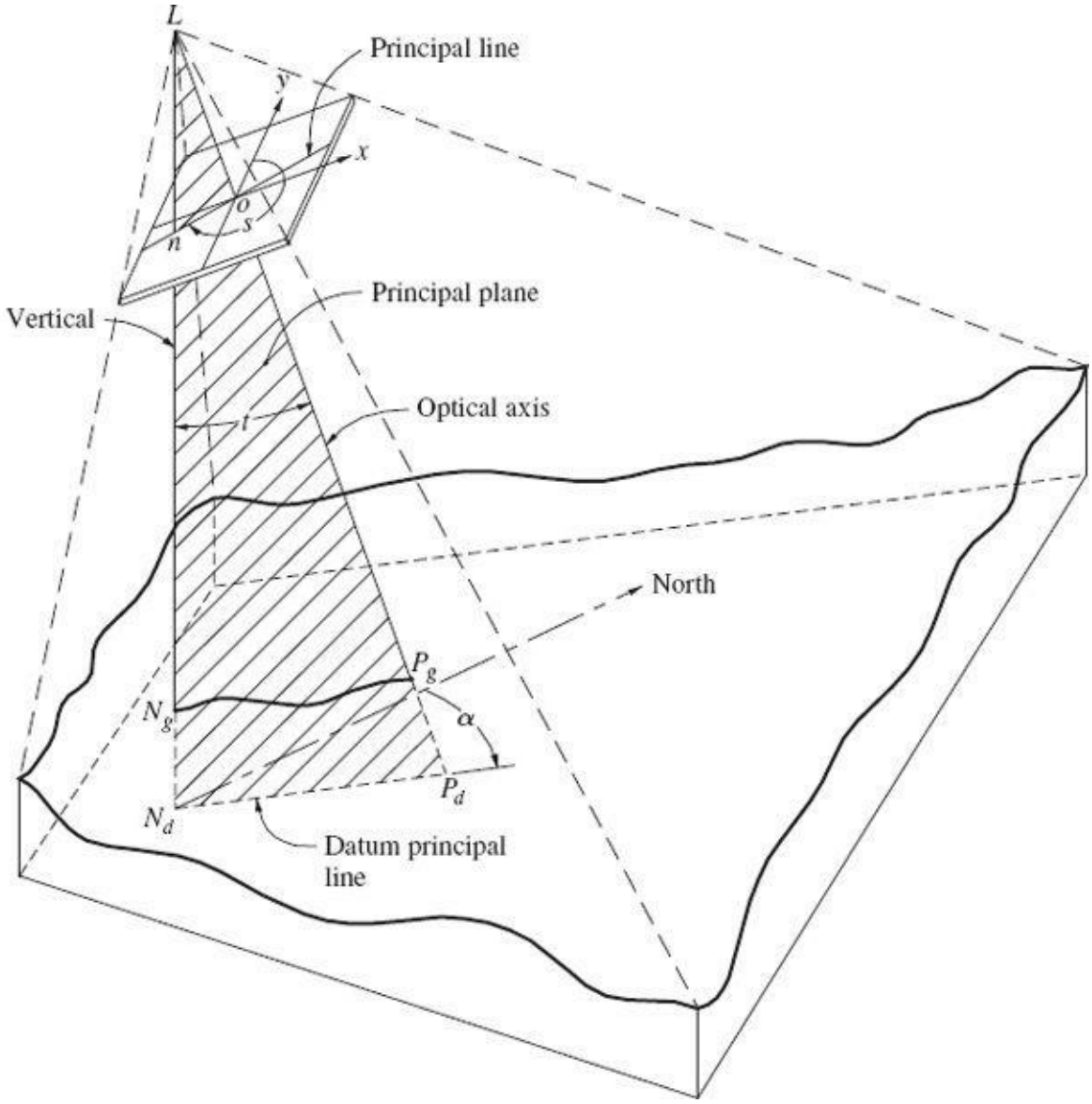


FIGURE 10-4 Geometry of a tilted photograph showing tilt-swing-azimuth angular orientation.

Vertical plane *Lno* is called the *principal plane*. Its line of intersection with the plane of the photograph occurs along line *no*, which is called the *principal line*. The position of the principal line on the photo with respect to the reference fiducial axis system is given by *s*, the *swing* angle. *Swing* is defined as the clockwise angle measured in the plane of the photograph from the positive *y* axis to the downward or nadir end of the principal line, as shown in Fig. 10-4. The swing angle gives the direction of tilt on the photo. Its value can be anywhere between 0° and 360°, or alternatively, between –180° and 180°.

The third angular orientation parameter, α or *azimuth*, gives the orientation of the principal plane with respect to the ground reference axis system. Azimuth is the clockwise angle measured from the ground *Y* axis (usually north) to the *datum principal line*, $N_d P_d$. It is measured in the datum plane or in a plane parallel to the datum plane, and its value can be anywhere between 0° and 360°, or alternatively, between –180° and 180°. The three angles of tilt, swing, and azimuth completely define the angular orientation of a tilted photograph in space. If the tilt angle is zero, the photo is vertical. Thus a vertical photograph is simply a special case of the general tilted photograph. For a vertical photo, swing and azimuth are undefined, but arbitrary definitions can be provided (see Sec. D-10).

10-4 Auxiliary Tilted Photo Coordinate System

In the tilt-swing-azimuth system, certain computations require the use of an auxiliary $x'y'$ rectangular photographic coordinate system. This system, as shown in Fig. 10-5*a*, has its origin at the photographic nadir point n and its y' axis coincides with the principal line (positive in the direction from n to o). Positive x' is 90° clockwise from positive y'. In solving tilted photo problems in the tilt-swing-azimuth system, photo coordinates are usually first measured in the fiducial coordinate system described in Sec. 4-2 and then converted to the auxiliary system numerically.

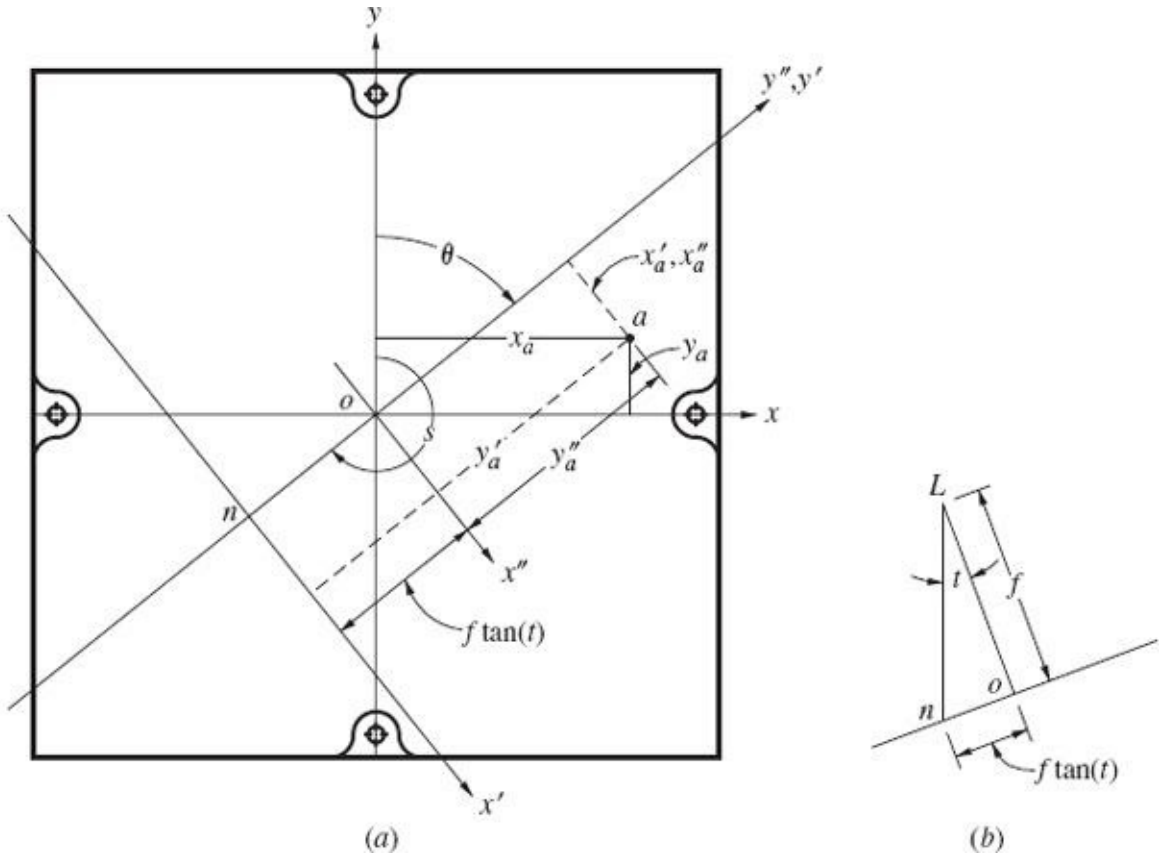


FIGURE 10-5 (*a*) Auxiliary x′y′ image coordinate system for a tilted photograph. (*b*) Principal plane of a tilted photo.

For any point in a tilted photo, the conversion from the xy fiducial system to the $x'y'$-tilted system requires (1) a rotation about the principal point through the angle θ and (2) a translation of origin from o to n. The rotation angle θ is defined as

$$\theta = s - 180° \tag{10-1}$$

The coordinates of image point a after rotation are x''_a and y''_a, as shown in Fig. 10-5*a*. These coordinates, calculated in the same fashion as the E'_c and N'_c coordinates of Eqs. (C-2), are specified by

$$x''_a = x_a \cos\theta - y_a \sin\theta$$

$$y''_a = x_a \sin\theta + y_a \cos\theta$$

Auxiliary coordinate y'_a is obtained by adding the translation distance on to y''_a. From Fig. 10-5*b*, which

is a side view of the principal plane, *on* is $f \tan t$. Since there is no x translation, $x'_a = x''_a$ and from the relationship between s and θ in Eq. (10-1), $\sin\theta = -\sin s$ and $\cos\theta = -\cos s$. Therefore the coordinates x'_a and y'_a of a point a in the required auxiliary coordinate system are

$$x'_a = -x_a \cos s + y_a \sin s$$

$$y'_a = -x_a \sin s - y_a \cos s + f \tan t \tag{10-2}$$

In Eqs. (10-2), x_a and y_a are the fiducial coordinates of point a, f is the camera focal length, and t and s are the tilt and swing angles, respectively.

10-5 Scale of a Tilted Photograph

In Chap. 6 it was shown that scale variations in a vertical photograph are the result of variations in object distance (distance from camera to ground). The shorter the object distance, the larger the scale, and vice versa. For vertical photos, variations in object distances were caused only by topographic relief. In a tilted photograph, relief variations also cause changes in scale, but scale in various parts of the photo is further affected by the magnitude and angular orientation of the tilt. Figure 10-6*a* portrays the principal plane of a tilted photograph taken over a square grid on approximately flat ground. Figure 10-6*b* illustrates the appearance of the grid on the resulting tilted photograph. Due to tilt, object distance *LA*′ in Fig. 10-6*a* is less than object distance *LB*′, and hence a grid line near *A* would appear larger (at a greater scale) than a grid line near *B*. This is illustrated in Fig. 10-6*b*, where photo distance d_1 appears longer than photo distance d_2, yet both are the same length on the ground.

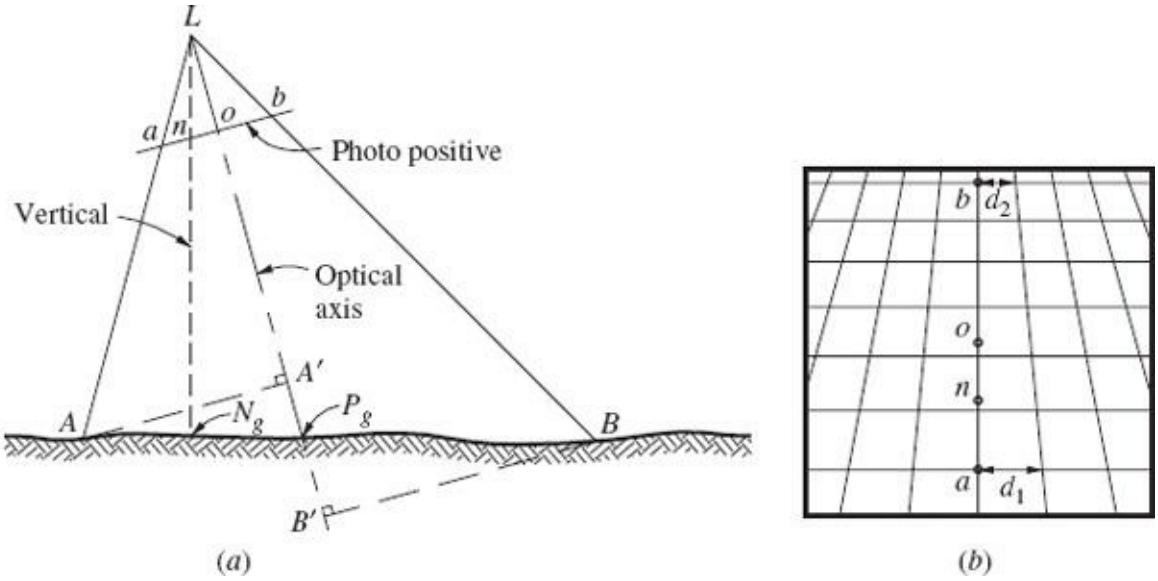


FIGURE 10-6 (*a*) Principal plane of a tilted photograph taken over approximately flat ground. (*b*) Image on the tilted photo of a square grid.

The scale at any point on a tilted photograph is readily calculated if tilt and swing for the photograph and the flying height of the photo and elevation of the point above datum are known. Figure 10-7 illustrates a tilted photo taken from a flying height *H* above datum; *Lo* is the camera focal

length. The image of object point *A* appears at *a* on the tilted photo, and its coordinates in the auxiliary tilted photo coordinate system are x'_a and y'_a. The elevation of object point *A* above datum is h_A. Object plane *AA′KK′* is a horizontal plane constructed at a distance h_A above datum. Image plane *aa′kk′* is also constructed horizontally. The scale relationship between the two parallel planes is the scale of the tilted photograph at point *a* because the image plane contains image point *a* and the object plane contains object point *A*. The scale relationship is the ratio of photo distance *aa′* to ground distance *AA′* and may be derived from similar triangles *La′a* and *LA′A*, and *Lka′* and *LKA′* as follows:

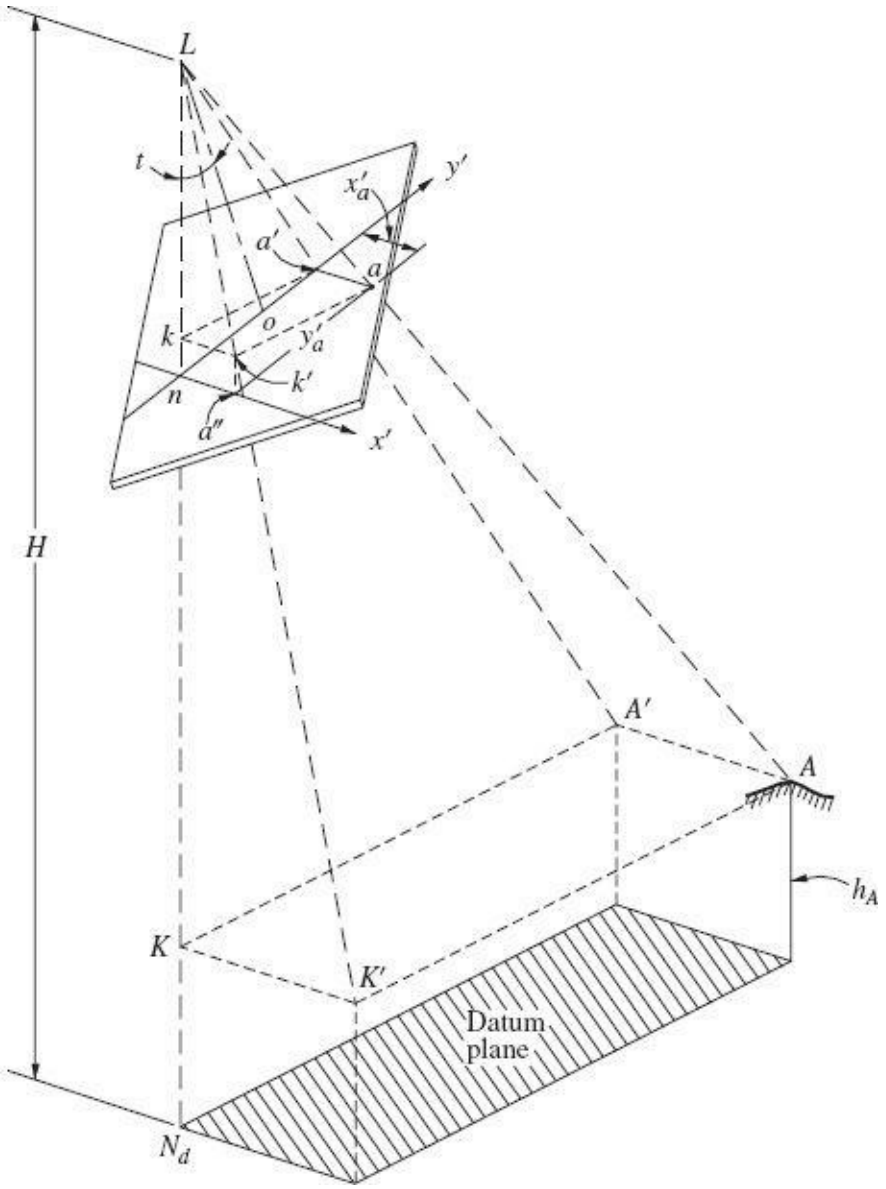


FIGURE 10-7 Scale of a tilted photograph.

$$S_a = \frac{aa'}{AA'} = \frac{La'}{LA'} = \frac{Lk}{LK} \tag{a}$$

but

$$Lk = Ln - kn = \frac{f}{\cos t} - y'_a \sin t$$

also

$$LK = H - hA$$

Substituting *Lk* and *LK* into Eq. (*a*) and dropping subscripts yields

$$S = \frac{f/\cos t - y' \sin t}{H - h} \tag{10-3}$$

In Eq. (10-3), S is the scale on a tilted photograph for any point whose elevation is h above datum. Flying height above datum for the photo is H; f is the camera focal length; and y' is the coordinate of the point in the auxiliary system calculated by Eqs. (10-2). If the units of f and y' are millimeters and if H and h are meters, then the scale ratio is obtained in millimeters per meter. To obtain a dimensionless ratio, the right side of Eq. (10-3) must be multiplied by 1 m/1000 mm in that case. Examination of Eq. (10-3) shows that scale increases with increasing terrain elevation. If the photo is taken over level ground, then h is constant but scale still varies throughout the photograph with variations in y'.

Example 10-1
A tilted photo is taken with a 152.4-mm-focal-length camera from a flying height of 2266 m above datum. Tilt and swing are 2.53° and 218.20°, respectively. Point *A* has an elevation of 437 m above datum, and its image coordinates with respect to the fiducial axis system are $x_a = -72.4$ mm and 87.1 mm. What is the scale at point *a?*

Solution By the second equation of Eqs. (10-2),

$$y'_a = 72.4 \sin 218.2° - 87.1 \cos 218.20° + 152.4 \tan 2.53° = 30.4 \text{ mm}$$

By Eq. (10-3),

$$S_a = \frac{152.4/\cos 2.53° - 30.4 \sin 2.53°}{2266 - 437}\left(\frac{1 \text{ m}}{1000 \text{ mm}}\right) = 1 : 12\bar{1}\bar{0}0 \ (1 \text{ in}/10\bar{1}\bar{0} \text{ ft})$$

▲

10-6 Relief Displacement on a Tilted Photograph

Image displacements on tilted photographs caused by topographic relief occur much the same as they do on vertical photos, except that relief displacements on tilted photographs occur along radial lines from the nadir point. (Relief displacements on a truly vertical photograph are also radial from the nadir point, but in that special case the nadir point coincides with the principal point.)

On a tilted photograph, image displacements due to relief vary in magnitude depending upon flying height, height of object, amount of tilt, and location of the image in the photograph. Relief displacement is zero for images at the nadir point and increases with increased radial distances from the nadir.

As defined in Sec. 10-1, tilted photos are those that were intended to be vertical, but contain small unavoidable amounts of tilt. In practice, tilted photos are therefore so nearly vertical that their

nadir points are generally very close to their principal points. Even for a photograph containing 3° of tilt taken with a 152-mm-focal-length camera, distance *on* is only about 8 mm. Relief displacements on tilted photos may therefore be calculated with satisfactory accuracy using Eq. (6-11), which applies to a vertical photograph. When this equation is used, radial distances r are measured from the principal point, even though theoretically, they should be measured from the nadir point.

10-7 Determining the Angle of Inclination of the Camera Axis in Oblique Photography

Oblique photos, in which the camera axis is inclined either up or down from horizontal at the time of exposure, represent the typical case in terrestrial photogrammetry. Since the base configuration for terrestrial photos is a horizontal pose, it is useful to characterize its orientation by the angular deviation from horizontal. Figure 10-8 illustrates an oblique terrestrial photo taken with the camera axis inclined at an angle θ (not to be confused with θ defined in Sec. 10-4) from horizontal. In this case the axis is inclined downward, and θ is called a *depression* angle. If the camera axis were inclined upward, θ would be an *elevation* angle. The angle of inclination of the camera axis is an important variable for certain elementary methods of determining object space positions of points whose images appear on overlapping terrestrial photos. With some terrestrial cameras the angle of inclination can be set or measured so that it becomes a known quantity. If it is unknown for a particular photo or photos, methods are available for computing its value.

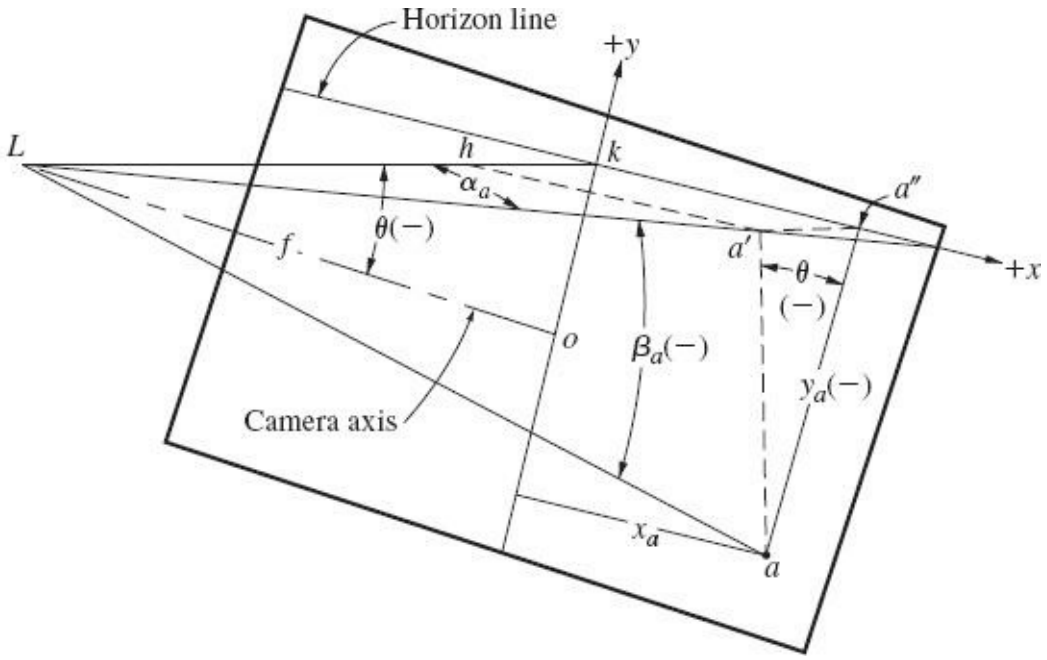


FIGURE 10-8 Horizontal and vertical angles from measurements on an oblique terrestrial photo.

One elementary method of determining the angle of inclination of the camera axis of a terrestrial photo relies on the fundamental principles of perspective geometry. If a photograph contains images of linear features which are horizontal or vertical, the horizon or nadir can be located through graphical construction. Figure 10-9 illustrates an oblique terrestrial photo of two buildings at a street corner. In this figure, dashed lines are extended from the tops and bottoms of the windows (horizontal parallel lines) to their intersections at vanishing points $v1$ and $v2$, which when connected define the horizon. Also shown are dashed lines extending from the vertical building edges to their intersection

at the nadir point *n*. The line from *n* through the principal point *o* intersects the horizon at a right angle at point *k*.

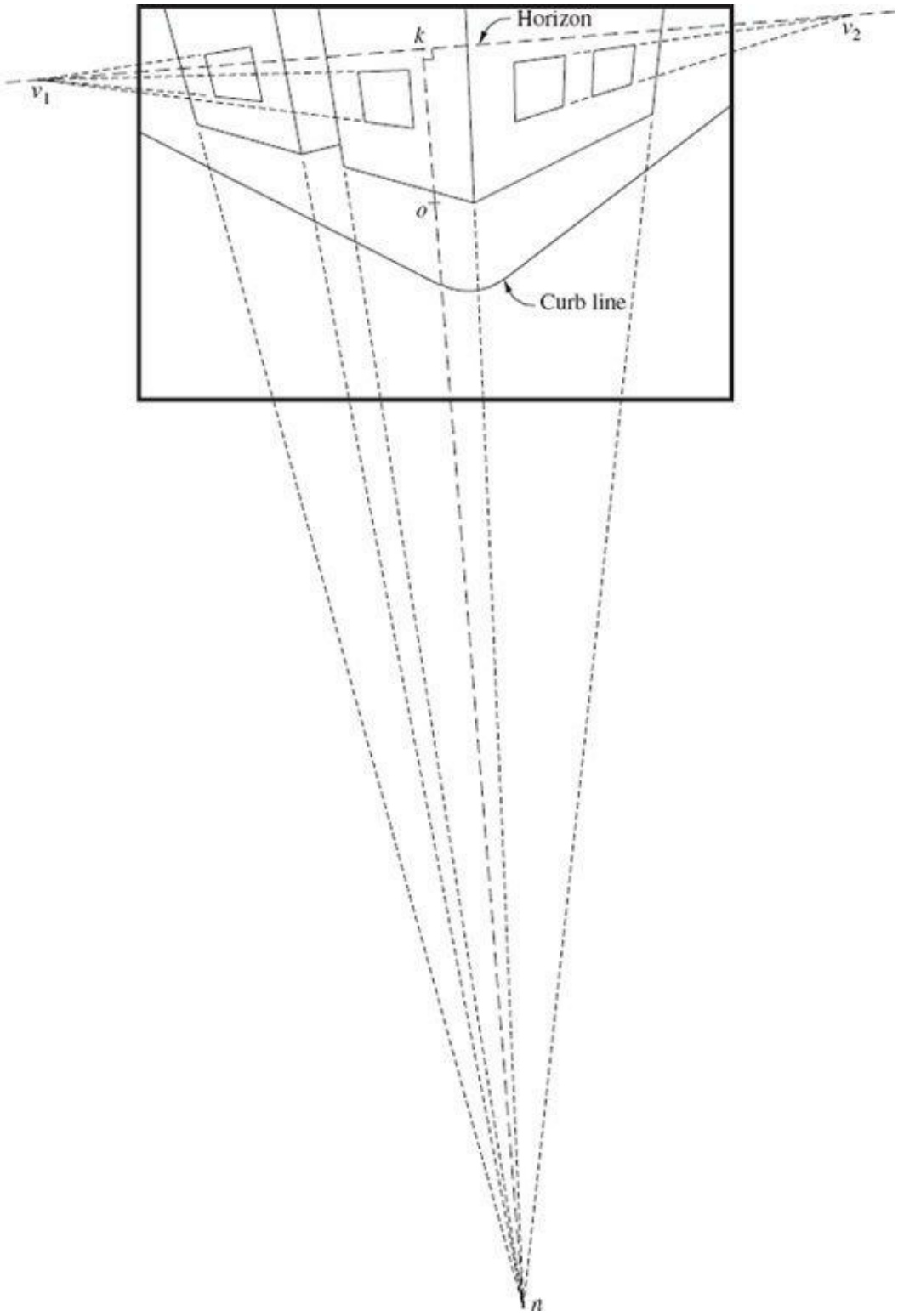


FIGURE 10-9 Location of horizon and nadir on an oblique photograph.

A reference photo coordinate system can be established for the oblique photo of Fig. 10-9 from the preceding graphical construction. Figure 10-10*a* shows the *x* and *y* axes of the oblique photo coordinate system with its origin at point *k*. Note that the *x* axis coincides with the horizon. Figure 10-10*b* shows a profile view of the principal plane *Lkn*. The optical axis *Lo* is inclined downward at a depression angle θ. To conform with usual sign convention, depression angles are considered negative

in algebraic sign. If the camera axis is inclined upward, θ is considered positive. Angle θ can be determined by either of two approaches. The first approach requires that the horizon be located by vanishing points. Then line *ko* is drawn at a right angle to the horizon through point *o* (the principal point of the photo) to define the *y* axis. Distance *ko* is then measured, and angle θ is computed from the following (see Fig. 10-10*b*):

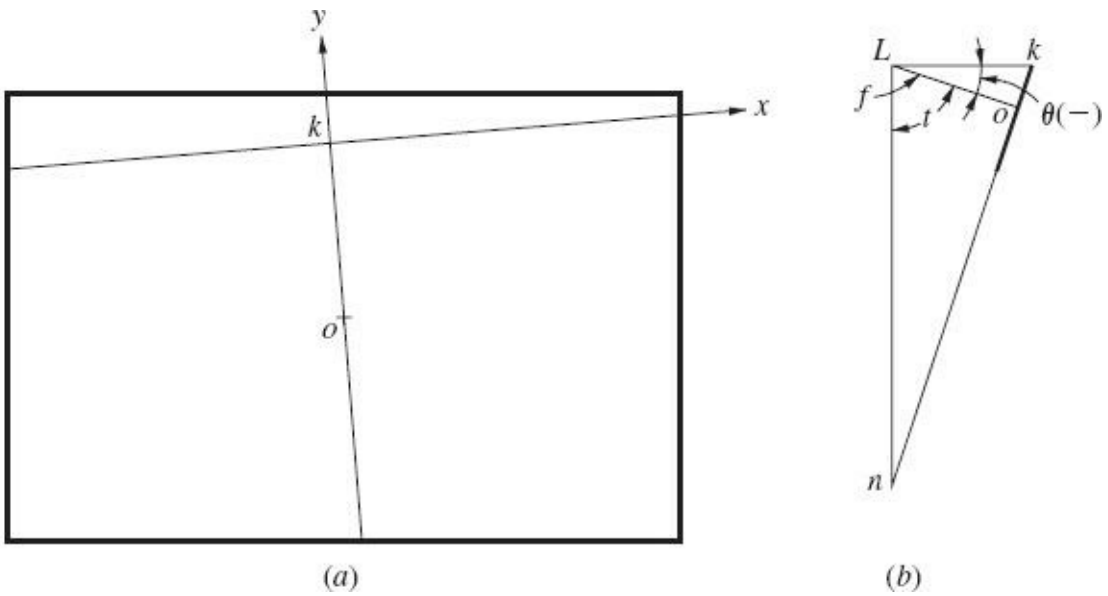


FIGURE 10-10 (*a*) Oblique photo coordinate axis system. (*b*) Side view of principal plane showing depression angle θ and tilt angle *t*.

$$\theta = \tan^{-1}\left(\frac{y_0}{f}\right) \tag{10-4}$$

In Eq. (10-4), θ is the depression (or elevation) angle, y_0 is the *y* coordinate of the principal point in the oblique photo coordinate system, and *f* is the camera focal length. In this equation, the correct algebraic sign must be applied to y_0 so that angle θ, in turn, will have the correct algebraic sign. It is also necessary to use an appropriate value for the focal length *f*. If the graphical analysis is being made on an enlarged photographic print, the focal length of the camera must be correspondingly enlarged. The second approach to determining angle θ is based on the location of the nadir. After the nadir has been determined, distance *on* is measured and the tilt angle *t* computed by the following:

$$t = \tan^{-1}\left(\frac{on}{f}\right) \tag{10-5}$$

Once the tilt angle has been determined, angle θ can be computed from

$$\theta = t - 90^\circ \tag{10-6}$$

If a photograph is taken in which the optical axis is pointed upward from the horizon, angle θ is an elevation angle, and vertical parallel lines will intersect at the zenith. In this case, the distance *oz* from the principal point to the zenith point is measured, and angle θ can be computed by

$$\theta = 90° - \tan^{-1}\left(\frac{oz}{f}\right) \tag{10-7}$$

As discussed in Sec. 19-8, analytical methods can also be used for determining the angle of inclination of the camera axes for terrestrial photos, and these methods provide the highest levels of accuracy.

10-8 Computing Horizontal and Vertical Angles from Oblique Photos

Once angle θ has been determined, horizontal and vertical angles can be computed for points whose images appear in the photograph. Refer to Fig. 10-8 which shows the geometry of an oblique terrestrial photograph. In this figure, L is the exposure station, and Lo is f, the camera focal length; Lk is a horizontal line intersecting the photo at k. The x axis coincides with the horizon line, and the y axis passes through the principal point o, perpendicular to the x axis at point k. Point a is an image point, and aa' is a vertical line, with a' being located in the horizontal plane of the exposure station and horizon line. Horizontal angle α_a between the vertical plane, ($L\ a'\ a$), containing image point a and the vertical plane, Lko, containing the camera axis is

$$\alpha_a = \tan^{-1}\left(\frac{ha'}{Lk - hk}\right) = \tan^{-1}\left(\frac{x_a}{f\sec\theta - y_a\sin\theta}\right) \tag{10-8}$$

In Eq. (10-8) note that correct algebraic signs must be applied to x_a, y_a, and θ. Algebraic signs of α angles are positive if they are clockwise from the optical axis and negative if they are counterclockwise. After the horizontal angle α_a has been determined, vertical angle β_a to image point a can be calculated from the following equation:

$$\beta_a = \tan^{-1}\left(\frac{aa'}{La'}\right) = \tan^{-1}\left(\frac{aa'}{(Lk - hk)\sec\alpha_a}\right)$$

$$= \tan^{-1}\left(\frac{y_a\cos\theta}{(f\sec\theta - y_a\sin\theta)\sec\alpha_a}\right) \tag{10-9}$$

The algebraic signs of β angles are automatically obtained from the signs of the y coordinates used in Eq. (10-9).

Example 10-2

An oblique terrestrial photo was exposed with the camera axis depressed at an angle $\theta = -14.3°$. The camera focal length was 60.00 mm. Compute the horizontal and vertical angles to an object point A whose image has photo coordinates $x_a = 27.4$ mm and $y_a = 46.2$ mm, measured with respect to the oblique photo coordinate axes.

Solution. From Eq. (10-8),

$$\alpha_a = \tan^{-1}\left(\frac{27.4}{60.00\sec(-14.3°)-(-46.2)\sin(-14.3°)}\right) = 28.5°$$

From Eq. (10-9),

$$\beta_a = \tan^{-1}\left(\frac{-46.2\cos(-14.3°)}{[60.00\sec(-14.3°)-(-46.2)\sin(-14.3°)](\sec 28.5°)}\right) = -37.9°$$

▲

10-9 Angular Orientation in Omega-Phi-Kappa

As previously stated, besides the tilt-swing-azimuth system, angular orientation of a photograph can be expressed in terms of three rotation angles, *omega*, *phi*, and *kappa*. These three angles uniquely define the angular relationships between the three axes of the photo (image) coordinate system and the three axes of the ground (object) coordinate system.

Figure 10-11 illustrates a tilted photo in space. The ground coordinate system is *XYZ*. The tilted photo image coordinate system is *xyz* (shown dashed), and its origin is at exposure station *L*. Consider an image coordinate system *x′y′z′* with origin also at *L* and with its respective axes mutually parallel to the axes of the ground coordinate system, as shown in Fig. 10-11. As a result of three sequential rotations through the angles of omega, phi, and kappa, the *x′y′z′* axis system can be made to coincide with the photographic *xyz* system. Each of the rotation angles omega, phi, and kappa is considered positive if counterclockwise when viewed from the positive end of the rotation axis.

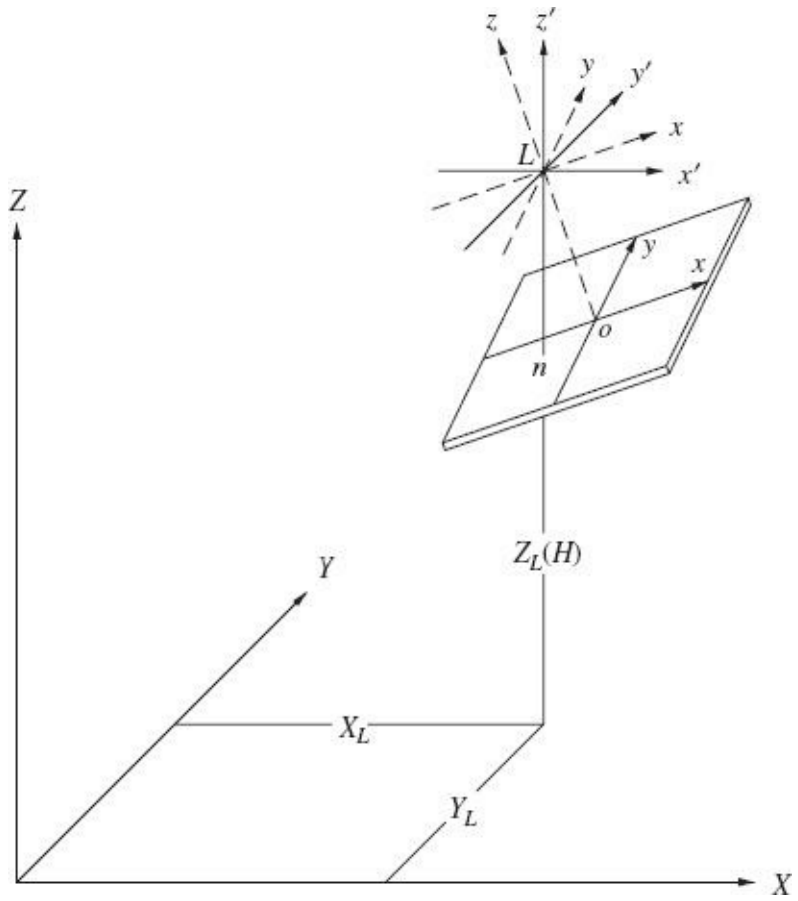


FIGURE 10-11 Orientation of a photograph in the omega-phi-kappa system.

The sequence of the three rotations is illustrated in Fig. 10-12. The first rotation, as illustrated in Fig. 10-12*a*, is about the x' axis through an angle omega. This first rotation creates a new axis system $x_1y_1z_1$. The second rotation phi is about the y_1 axis, as illustrated in Fig. 10-12*b*. As a result of the phi rotation a new axis system $x_2y_2z_2$ is created. As illustrated in Fig. 10-12*c*, the third and final rotation is about the z_2 axis through the angle kappa. This third rotation creates the *xyz* coordinate system which is the photographic image system. Equations that express these three rotations are developed in Sec. C-7.

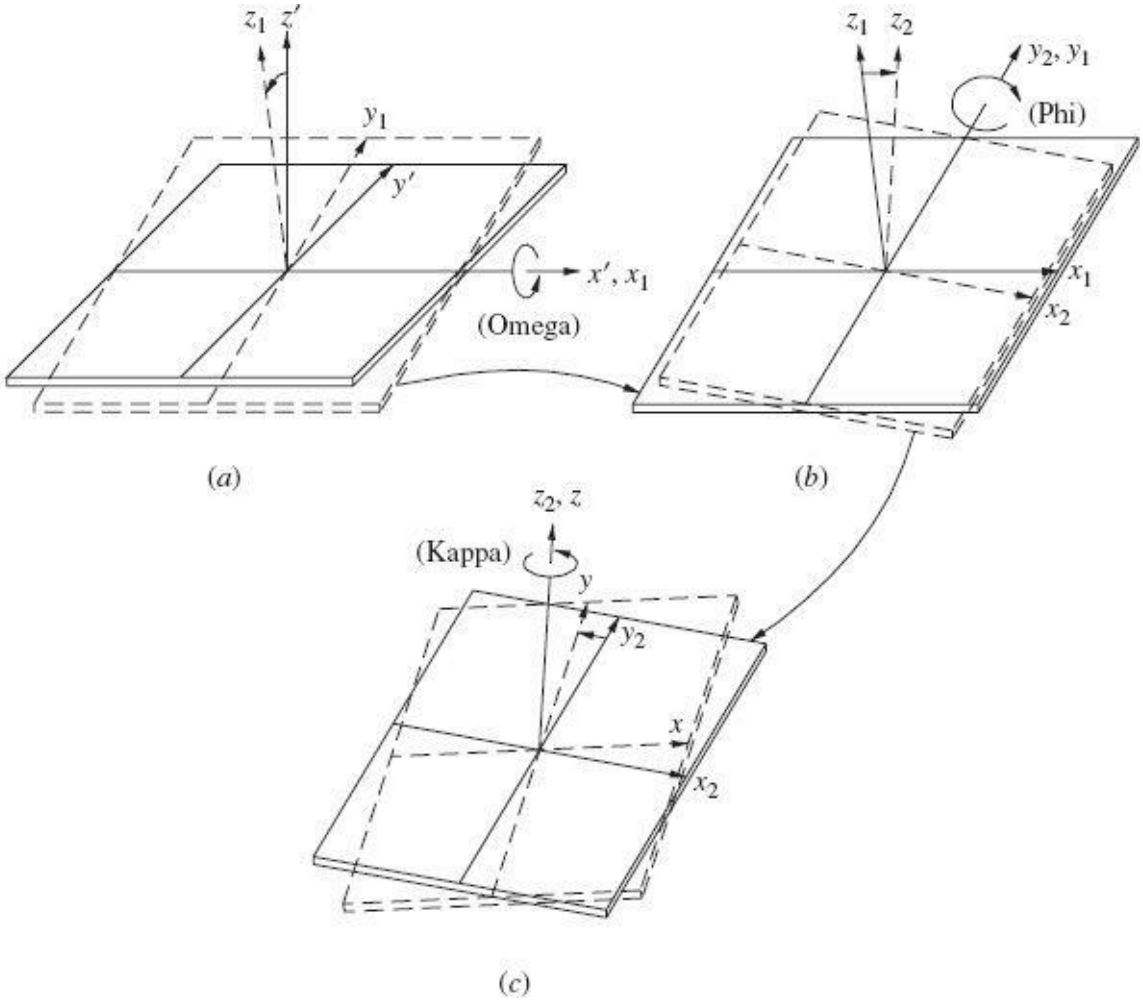


FIGURE 10-12 (*a*) Rotation about the x' axis through angle omega. (*b*) Rotation about the y_1 axis through angle phi. (*c*) Rotation about the z_2 axis through angle kappa.

For any photo there exist a unique set of angles omega, phi, and kappa which explicitly define the angular orientation of the photograph with respect to the reference ground coordinate system, provided appropriate ranges are maintained (see Sec. D-10). These three angles are related to the previously described tilt, swing, and azimuth angles; and if either set of three orientation angles is known for any photo, the other three can be determined as described in Sec. D-10. In the omega-phi-kappa system, as with the tilt-swing-azimuth system, the space position of any photo is given by the exposure station coordinates X_L, Y_L, and Z_L (or H).

10-10 Determining the Elements of Exterior Orientation

Many different methods, both graphical and numerical, have been devised to determine the six elements of exterior orientation of a single photograph. In general all methods require photographic images of at least three control points whose *X*, *Y*, and *Z* ground coordinates are known. Also, the calibrated focal length of the camera must be known. Although there are a number of methods to find these elements (see Chap. 17), the most commonly used is *space resection by collinearity*.

The method of space resection by collinearity is a purely numerical method which simultaneously yields all six elements of exterior orientation. Normally the angular values of omega, phi, and kappa are obtained in the solution, although the method is versatile and tilt, swing, and azimuth could also be obtained. Space resection by collinearity permits the use of redundant amounts of ground control; hence least squares computational techniques can be used to determine most probable values for the six elements. This method is a rather lengthy procedure requiring an iterative solution of nonlinear equations, but when programmed for a computer, a solution is easily obtained. Sometimes, when rapid computation is needed, a linear model is used, although with less precise results (see Chap. 17).

Space resection by collinearity involves formulating the so-called *collinearity equations* for a number of control points whose *X*, *Y*, and *Z* ground coordinates are known and whose images appear in the photo. The equations are then solved for the six unknown elements of exterior orientation which appear in them. The collinearity equations express the condition that for any given photograph the exposure station, any object point and its corresponding image all lie on a straight line. Figure D-1 illustrates the collinearity condition for exposure station *L*, image point *a*, and object point *A*. The details involved in the numerical solution of space resection by collinearity are presented in Chap. 11.

10-11 Rectification of Tilted Photographs

Rectification is the process of making equivalent vertical photographs from tilted photo negatives. The resulting equivalent vertical photos are called *rectified photographs*. Rectified photos theoretically are truly vertical photos, and as such they are free from displacements of images due to tilt. They do, however, still contain image displacements and scale variations due to topographic relief. These relief displacements and scale variations can also be removed in a process called *differential rectification* or *orthorectification*, and the resulting products are then called *orthophotos*. Orthophotos are often preferred over rectified photos because of their superior geometric quality. Nevertheless, rectified photos are still quite popular because they do make very good map substitutes where terrain variations are moderate. This chapter discusses only the methods of rectification for the removal of the effects of tilt. Orthophoto production is described in Chap. 13.

Rectification is generally performed by any of three methods: analytically, optically-mechanically, and digital. Analytical rectification has the disadvantage that it can be applied only to individual discrete points, i.e., points that can be specifically identified so that their locations within the tilted photo can be measured. The resulting rectified photos produced by analytical methods are not really photos at all since they are not composed of photo images. Rather, they are plots of individual points in their rectified locations. The optical-mechanical and digital methods produce an actual picture in which the images of the tilted photo have been transformed to their rectified locations. Thus, the products of these two methods can be used in the production of photomaps and mosaics. In any of these rectification procedures, the rectified photos can be simultaneously *ratioed*; i.e., their average scales can be brought to some desired value different from that of the original photo. This is particularly advantageous if rectified photos are being made for the purpose of constructing a

controlled mosaic, since all photos in the strip of block can be brought to a common scale. Thus, the resulting mosaic will have a more uniform scale throughout.

The fundamental geometry of rectification is illustrated in Fig. 10-13. This figure shows a side view of the principal plane of a tilted photo. When the exposure was made, the negative plane made an angle t with the datum plane. Rays from A and B were imaged at a' and b', respectively, on the negative, and their corresponding locations on the tilted photo are at a and b. The plane of an *equivalent vertical photo* is shown parallel to the datum plane and passing through i, the *isocenter* of the tilted photo.[1] The plane of a ratioed rectified photo is also shown. It is likewise parallel to the datum plane but exists at a level other than that of the equivalent vertical photo plane.

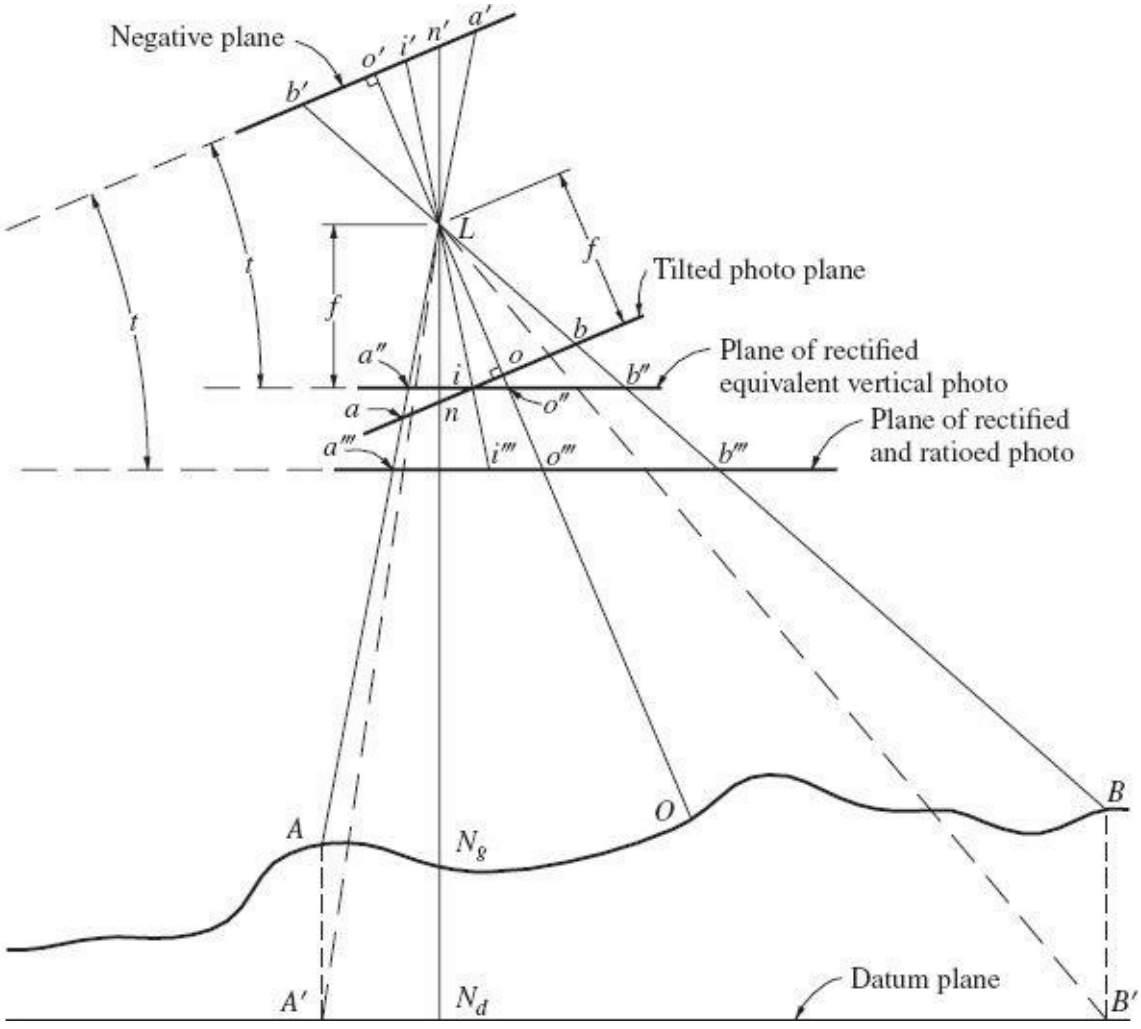


FIGURE 10-13 Principal plane of a tilted photograph showing the basic geometry of rectification.

Methods of projecting points such as a and b to either a' and b' or a''' and b''' are subjects of the following sections of this chapter. Figure 10-13 also illustrates, by virtue of lines LA' and LB', that although tilt displacements are removed in rectification, displacements due to relief are still present.

10-12 Correction for Relief of Ground Control Points Used in Rectification

Since rectified and ratioed aerial photos retain the effects of relief, ground control points used in rectification must be adjusted slightly to accommodate their relief displacements. Conceptually, the

process involves plotting ground control points at the locations that they will occupy in the rectified and ratioed photo. To this end, the positional displacements due to relief of the control points must be computed and applied to their horizontal positions in a radial direction from the exposure station so that they will line up with points in the rectified photo. This procedure requires that the coordinates X_L, Y_L, and Z_L (or H) of the exposure station (which can be computed by space resection) and the X, Y, and Z (or h) coordinates for each ground control point be known. In addition, when the rectified photos are to be ratioed as well, it is convenient to select a plane in object space, at a specified elevation, to which the scale of the ratioed photo will be related. Generally, the elevation of this plane will be chosen as the elevation of average terrain, h_{avg}.

The procedure is illustrated in Fig. 10-14, which shows the horizontal position of the exposure station, L, denoted as a cross and the unadjusted horizontal positions of four ground control points, A, B, C, and D, denoted as circles. Also illustrated are radial distances r' from L to each of the points, as well as the relief displacements, d, from the location in the plane of average terrain to the photo locations, A', B', C', and D', of the control points, denoted as triangles. Since the elevations of these control points may be higher or lower than the average terrain, the relief displacements may be either outward (control point higher than average terrain) or inward (control point lower than average terrain). Note that in Fig. 10-14, the elevations of points B and D are less than average terrain elevation and the elevations of points A and C are greater.

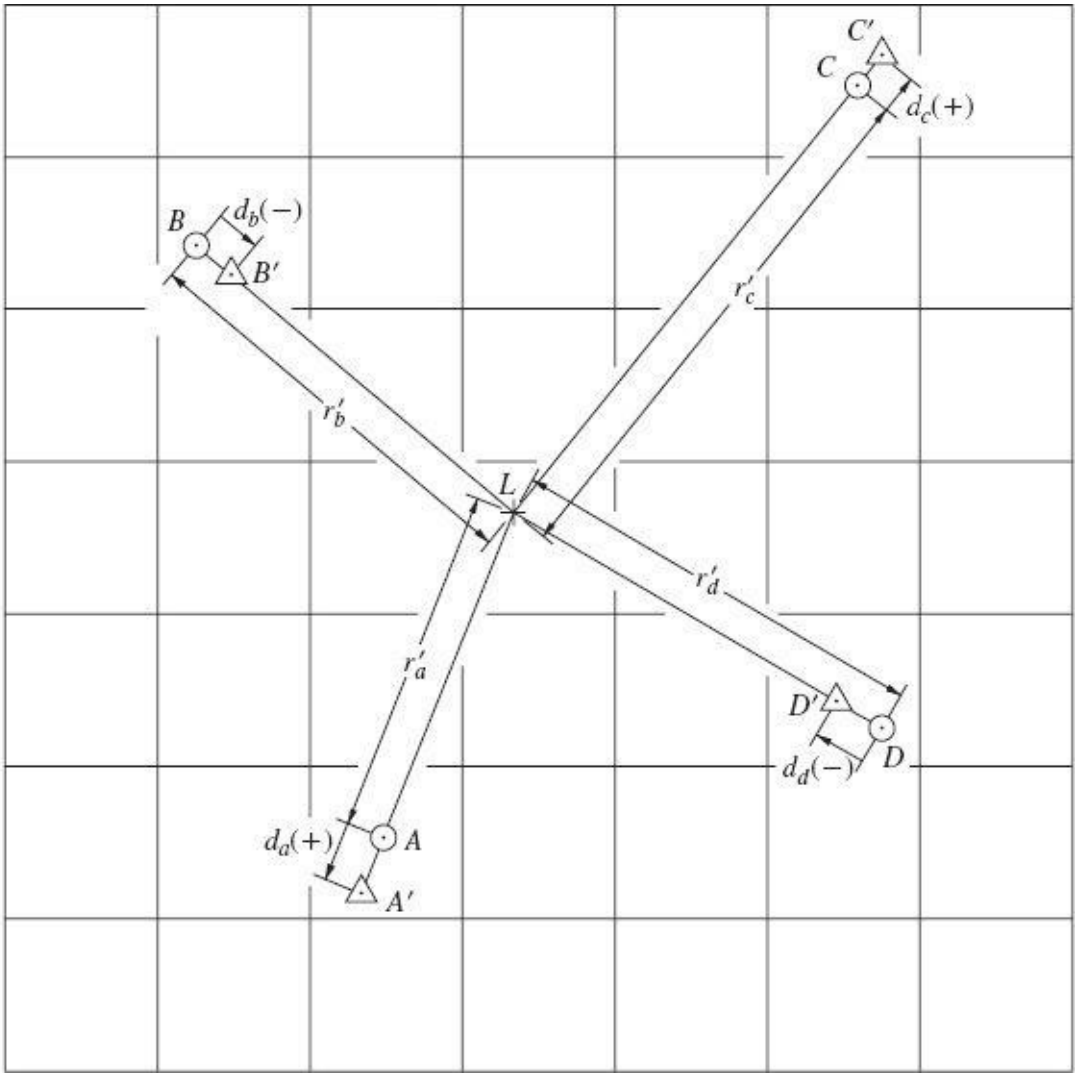


FIGURE 10-14 Plot of control points for rectification showing corrections made for relief displacements.

Determination of the coordinates of a displaced point (triangle) involves several steps. Initially, the value of r' is computed by Eq. (10-10) using the point's horizontal coordinates X and Y, and the horizontal coordinates of the exposure station, X_L and Y_L.

$$r' = \sqrt{(X - X_L)^2 + (Y - Y_L)^2} \tag{10-10}$$

Next, the relief displacement, d, is calculated by Eq. (10-11), which is a variation of Eq. (6-11):

$$d = \frac{r'(h - h_{avg})}{H - h} \tag{10-11}$$

In Eq. (10-11), d is the relief displacement, r' is the radial distance [computed by Eq. (10-10)], h the height of the control point above datum, h_{avg} the average terrain height in the tilted photo (also the height above datum of the ratioed photo plane), and H the flying height above datum for the tilted photo. The units of all terms in the equation are those of the object space coordinate system, i.e., either meters or feet. Once the displacement d has been computed, the radial distance r to the displaced (image) location of the point can be computed by the following equation (be careful to use the proper algebraic sign of d).

$$r = r' + d \tag{10-12}$$

Next, the azimuth, α, from the exposure location to the control point can be computed by

$$\alpha = \tan^{-1}\left(\frac{X - X_L}{Y - Y_L}\right) \tag{10-13}$$

In Eq. (10-13), it is necessary to use the full circle inverse tangent function so that the entire range of azimuth can be determined. Finally, the X' and Y' coordinates of the displaced (image) point are computed by

$$X' = X_L + r\sin\alpha \tag{10-14}$$

$$Y' = Y_L + r\cos\alpha \tag{10-15}$$

The resulting coordinates, X' and Y', are appropriate for use in any of the methods of rectification.

10-13 Analytical Rectification

There are several methods available for performing analytical (numerical) rectification. Each of the analytical methods performs rectification point by point, and each requires that sufficient ground control appear in the tilted photo. Basic input required for the numerical methods, in addition to

ground coordinates of control points, is x and y photo coordinates of all control points plus those of the points to be rectified. These are normally measured on a comparator. Due to the lengthy calculations required, numerical rectification is generally performed through the use of a computer program.

Of the available methods of analytical rectification, the one that uses the two-dimensional projective transformation is the most convenient and is the only method that will be discussed here. The transformation equations are developed in Sec. C-9 and are repeated here as Eqs. (10-16), for convenience:

$$X = \frac{a_1 x + b_1 y + c_1}{a_3 x + b_3 y + 1}$$

$$Y = \frac{a_2 x + b_2 y + c_2}{a_3 x + b_3 y + 1} \qquad (10\text{-}16)$$

In Eqs. (10-16), X and Y are ground coordinates, x and y are photo coordinates (in the fiducial axis system), and the a's, b's, and c's are the eight parameters of the transformation. The use of these equations to perform analytical rectification is a two-step process. First, a pair of Eq. (10-16) is written for each ground control point. Four control points will produce eight equations, so that a unique solution can be made for the eight unknown parameters. It is strongly recommended that more than four control points be used so that an improved solution can be arrived at by using least squares. An added benefit is that redundant measurements may provide the ability of detecting mistakes in the coordinates, something which is not afforded by the unique solution using four control points.

Once the eight parameters have been determined, the second step of the solution can be performed—that of solving Eqs. (10-16) for each point whose X and Y rectified coordinates are desired. After rectified coordinates have been computed in the ground coordinate system, they can be plotted at the scale desired for the rectified and ratioed photo.

This analytical method is only rigorous if the ground coordinates, X and Y, of Eqs. (10-16) have been modified for relief displacements as discussed in Sec. 10-12. If this is not done, a quasi-rectification results; although if the terrain is relatively flat and level, the errors will be minimal.

10-14 Optical-Mechanical Rectification

In practice, the optical-mechanical method is still being used, although digital methods have rapidly surpassed this approach. The optical-mechanical method relies on instruments called *rectifiers*. They produce rectified and ratioed photos through the photographic process of projection printing (see Sec. 2-12); thus they must be operated in a darkroom.

As illustrated in Fig. 10-15, the basic components of a rectifier consist of a lens, a light source with reflector, a stage for mounting the tilted photo negative, and an easel which holds the photographic emulsion upon which the rectified photo is exposed. The instrument is constructed with controls so that the *easel plane*, *lens plane* (plane perpendicular to the optical axis of the rectifier lens), and *negative plane* can be tilted with respect to each other, and controls that allow a shift to be applied to the negative plane. Thus provision is made for varying angles α and β of Fig. 10-15.

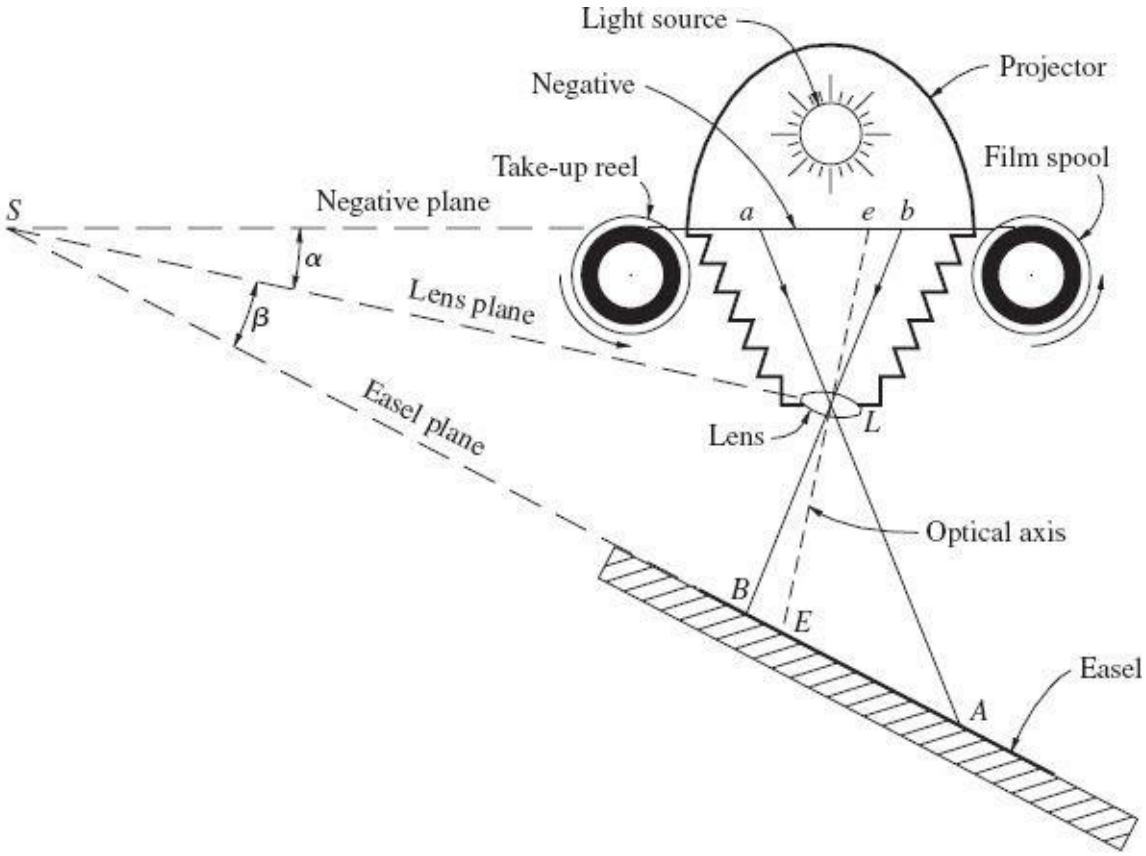


FIGURE 10-15 Schematic diagram of a tilting-lens optical-mechanical rectifier showing a side view of the principal plane.

So that rectified photos can be ratioed to varying scales, the rectifier must also have a magnification capability, and this is achieved by varying the projection distance (distance *LE* of Fig. 10-15 from the lens to the easel plane). To do this, however, and still maintain proper focus in accordance with Eq. (2-4), it is necessary to simultaneously vary the image distance (distance *Le* of Fig. 10-15 from the lens to the negative plane). The actual magnification that results along the axis of the rectifier lens is the ratio *LE*/(*Le*), but it varies elsewhere in the photo due to the variable scale of the tilted photograph.

From Fig. 10-15 it can be seen that in rectification, projection distances vary depending upon the locations of points on the photo. To achieve sharp focus for all images in spite of this, the *Scheimpflug condition* must be satisfied. The Scheimpflug condition states that, in projecting images through a lens, if the negative and easel planes are not parallel, the negative plane, lens plane, and easel plane must all intersect along a common line to satisfy the lens formula and achieve sharp focus for all images. Note that this condition is satisfied in Fig. 10-15 where these planes intersect at *S*.

10-15 Digital Rectification

Rectified photos can be produced by digital techniques that incorporate a photogrammetric scanner and computer processing. This procedure is a special case of the more general concept of georeferencing as presented in Chap. 9. The feature that distinguishes digital rectification from other georeferenced images is that rectification requires that a projective transformation be used to relate the image to the ground coordinate system whereas georeferencing often uses simpler transformations such as the two-dimensional conformal or the two-dimensional affine transformation.

Figure 10-16*a* shows a portion of a scanned aerial oblique. Note that images in the foreground

(near the bottom) are at a larger scale than those near the top. This photograph contains a generally rectangular street pattern, where the streets appear nonparallel due to the large amount of tilt. Figure 10-16*b* shows a digitally rectified version of the oblique photograph shown in Fig. 10-16*a*. Note that in this rectified image, the general street pattern is approximately rectangular, although displacements remain due to relief in the terrain as well as buildings.

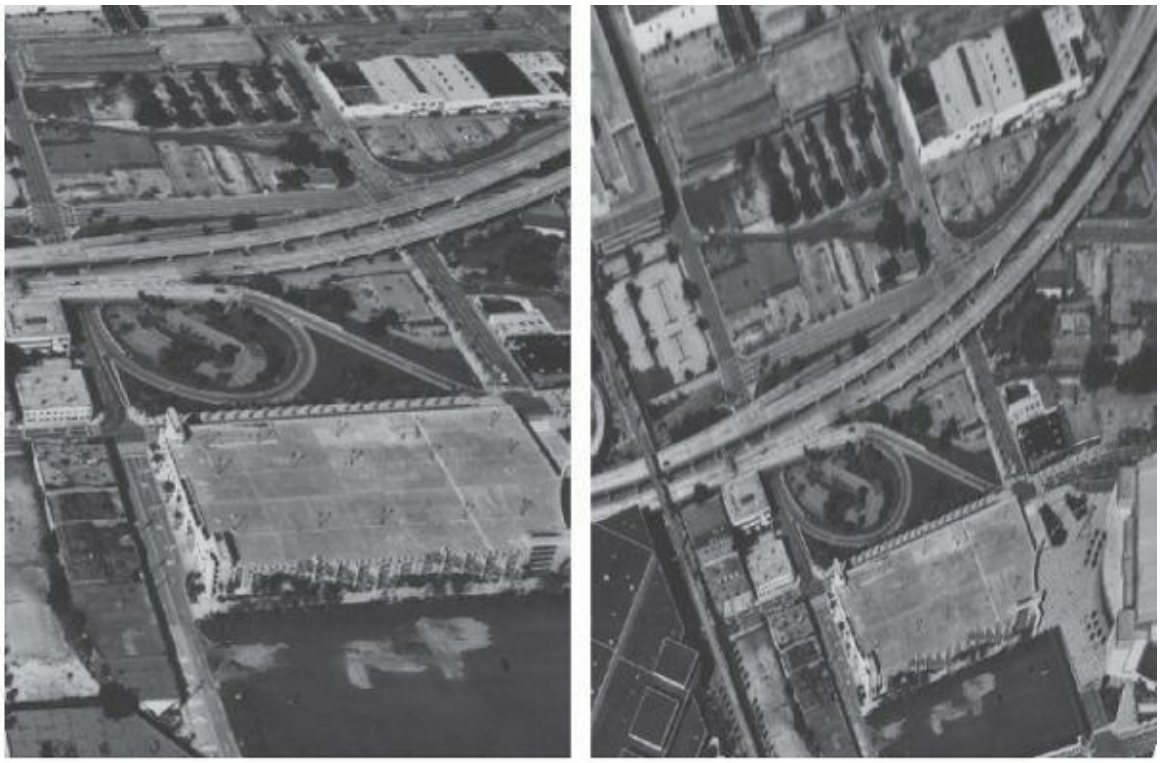


FIGURE 10-16 (*a*) Digital image of an oblique, nonrectified photograph. (*b*) Image after digital recification.

Three primary pieces of equipment needed for digital rectification are a digital image (either taken from a digital camera or a scanned photo), computer, and plotting device capable of producing digital image output. While a high-quality photogrammetric scanner is an expensive device, it is also highly versatile and can be used for many other digital-photogrammetric operations. Low-accuracy desktop scanners can also be used for digital rectification; however, the geometric accuracy of the product will be substantially lower. Current plotting devices, generally large-format ink jet printers, produce image output of good quality, although the geometric accuracy and image resolution may be slightly less than that of photographic rectifiers.

10-16 Atmospheric Refraction in Tilted Aerial Photographs

As mentioned in Sec. 4-12, atmospheric refraction in aerial photographs occurs radially from the nadir point. Now that parameters of tilted photographs have been described, a technique that corrects image coordinates in a tilted photo for the effects of atmospheric refraction is presented. In most practical situations, the assumption of vertical photography is sufficient when calculating atmospheric refraction. However for highest accuracy, when dealing with high altitude photography (e.g., flying height greater than 5000 m) and photographs with excessive tilts (e.g., tilts greater than 5°) should be considered when atmospheric refraction corrections are made.

Figure 10-17 illustrates a tilted photo containing image point a at coordinates x_a and y_a. The photographic nadir point n exists at coordinates x_n and y_n (in the fiducial axes coordinate system) which can be computed by the following equations. (Refer also to Fig. 10-2.)

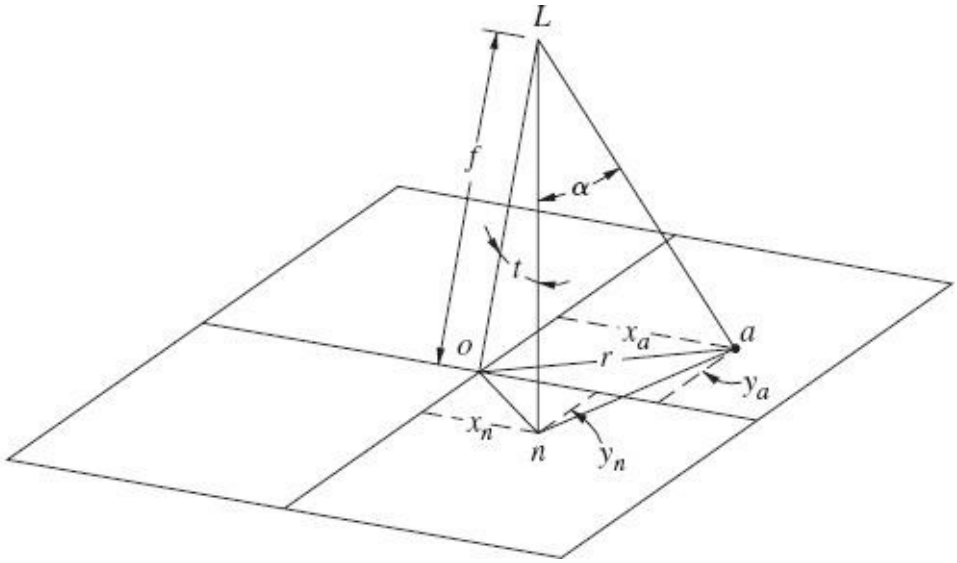


FIGURE 10-17 Diagram of a tilted photo with image point a whose coordinates will be corrected for atmospheric refraction.

$$x_n = on\sin s = f\tan t\sin s$$

$$y_n = on\cos s = f\tan t\cos s \tag{10-17}$$

The angle between the vertical line Ln and the incoming light ray through point a is designated as α. Angle α can be computed by the application of the law of cosines to triangle Lna. (Refer again to Figs. 10-5 and 10-17.)

$$(na)^2 = (Ln)^2 + (La)^2 - 2(Ln)(La)\cos\alpha$$

Rearranging gives

$$\alpha = \cos^{-1}\left[\frac{(Ln)^2 + (La)^2 - (na)^2}{2(Ln)(La)}\right] \tag{10-18}$$

where

$$na = \sqrt{(x_a - x_n)^2 + (y_a - y_n)^2} \tag{10-19}$$

$$Ln = \frac{f}{\cos t} \tag{10-20}$$

$$La = \sqrt{f^2 + r^2} = \sqrt{f^2 + x_a^2 + y_a^2} \tag{10-21}$$

After angle α has been determined, the refraction angle $\Delta\alpha$ can be calculated by Eq. (4-18) using a K value computed by Eq. (4-19).

Figure 10-18 shows triangle Lna along with the light ray La' which indicates the direction from the ground point that has been corrected for atmospheric refraction. From the figure, angle θ can be computed by the law of cosines as

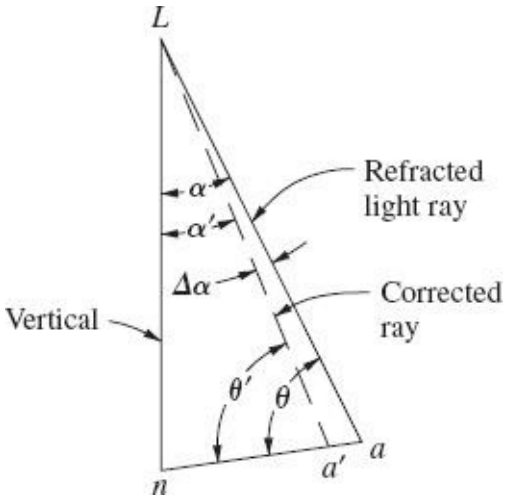


FIGURE 10-18 Vertical plane through tilted photo of Fig. 10-17 showing refracted light ray and corrected ray.

$$\theta = \cos^{-1}\left(\frac{(La)^2 + (na)^2 - (Ln)^2}{2(La)(na)}\right) \tag{10-22}$$

Angle θ' can then be computed by

$$\theta' = \theta + \Delta\alpha \tag{10-23}$$

Application of the law of sines to triangle Lna' gives

$$\frac{\sin\theta'}{Ln} = \frac{\sin\alpha'}{na'}$$

Solving for na' gives

$$na' = Ln\frac{\sin\alpha'}{\sin\theta'} \tag{10-24}$$

where

$$\alpha' = \alpha - \Delta\alpha \tag{10-25}$$

and Ln and θ' are computed by Eqs. (10-20) and (10-23), respectively. The displacement, aa' due to atmospheric refraction is then

$$aa' = na - na' \tag{10-26}$$

To compute the coordinate corrections δx and δy, displacement aa' must be applied in the direction of line an, as shown in Fig. 10-19. The direction of line an is specified by angle b as shown in the figure and can be computed by

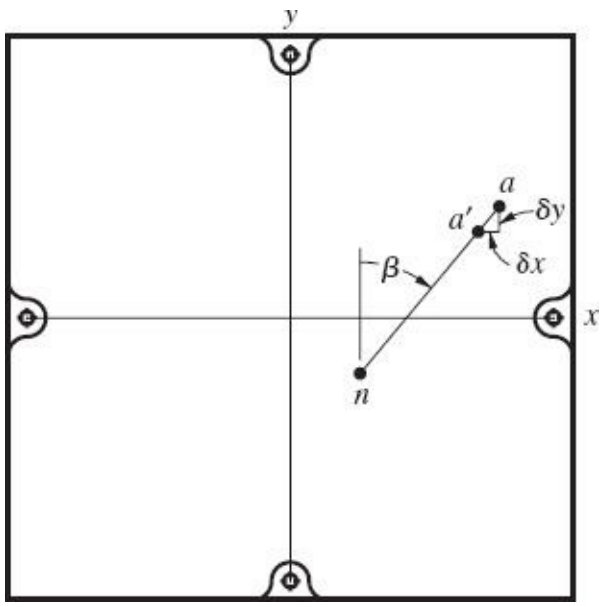


FIGURE 10-19 Photograph positions of image point a and its new location a' after correction for atmospheric refraction.

$$\beta = \tan^{-1}\left(\frac{x_a - x_n}{y_a - y_n}\right) \tag{10-27}$$

The corrections are then given by

$$\delta x = aa' \sin\beta$$

$$\delta y = aa' \cos\beta \tag{10-28}$$

Finally, the coordinates x'_a and y'_a for point a corrected for atmospheric refraction are computed by

$$x'_a = x_a - \delta x$$

$$y'_a = y_a - \delta y \tag{10-29}$$

Example 10-3

A tilted aerial photograph taken from a flying height of 4800 m above mean sea level contains the image *a* of object point *A* at fiducial coordinates x_a = –64.102 mm and y_a = 83.220 mm. The angles of tilt, swing, and azimuth for the photograph are 2.64°, 130.27°, and 5.15°, respectively. If the elevation of object point *A* is 140 m above mean sea level and the camera had a focal length of 152.794 mm, compute the x'_a and y'_a coordinates of the point, corrected for atmospheric refraction.

Solution Compute the fiducial coordinates of the nadir point *n* by Eqs. (10-17).

$$x_n = 152.794 \tan 2.64° \sin 130.27° = 5.376 \text{ mm}$$

$$y_n = 152.794 \tan 2.64° \cos 130.27° = -4.554 \text{ mm}$$

Compute the values of *na*, *Ln*, and *La* by Eqs. (10-19), (10-20), and (10-21), respectively.

$$na = \sqrt{(-64.102 - 5.376)^2 + (83.220 + 4.554)^2} = 111.944 \text{ mm}$$

$$Ln = \frac{152.794}{\cos 2.64°} = 152.956$$

$$La = \sqrt{(152.794)^2 + (-64.102)^2 + (83.220)^2} = 185.420 \text{ mm}$$

Compute angle α by Eq. (10-18).

$$\alpha = \cos^{-1}\left[\frac{(152.956)^2 + (185.420)^2 - (111.944)^2}{2(152.956)(185.420)}\right] = 37.0933°$$

Compute *K* by Eq. (4-19).

$$K = (7.4 \times 10^{-4})\ (4.8 - 0.14)\ \{1 - 0.02[2(4.8) - 0.14]\} = 0.0028°$$

Compute $\Delta\alpha$ by Eq. (4-18).

$$\Delta\alpha = 0.0028° \tan 37.0933° = 0.0021°$$

Compute θ by Eq. (10-22).

$$\theta = \cos^{-1}\frac{(185.420\ )^2 + (111.944)^2 - (152.956)^2}{2(185.420)(111.944)} = 55.4949°$$

Compute θ' and α' by Eqs. (10-23) and (10-25), respectively.

$$\theta' = 55.4949° + .0021° = 55.4970°$$

$$\alpha' = 37.0933° - .0021° = 37.0912°$$

Compute *na*′ by Eq. (10-24).

$$na' = 152.956\frac{\sin 37.0912^\circ}{\sin 55.4970^\circ} = 111.935 \text{ mm}$$

Compute aa' and β by Eqs. (10-26) and (10-27), respectively.

$$aa' = 111.944 - 111.935 = 0.008 \text{ mm}$$

$$\beta = \tan^{-1}\frac{-64.102 - 5.376}{83.220 - (-4.554)} = -38.3634^\circ$$

Compute δx and δy by Eqs. (10-28).

$$\delta x = 0.008 \sin 38.3634^\circ = -0.005 \text{ mm}$$

$$\delta y = 0.008 \cos 38.3634^\circ = 0.007 \text{ mm}$$

The corrected coordinates x'_a and y'_a are computed by Eqs. (10-29).

$$x'_a = -64.102 - (-0.005) = -64.097 \text{ mm}$$

$$y'_a = 83.220 - 0.007 = 83.213 \text{ mm} \quad \blacktriangle$$

Two final comments regarding refraction in a tilted photo are in order. First, the values of tilt and swing are not known until after an analytical solution is performed (see Chap. 11). However, photo coordinates are necessary in order to compute the analytical solution. Therefore refinement for atmospheric refraction in a tilted photograph must be performed in an iterative fashion. Since analytical photogrammetric solutions are generally iterative due to the nonlinear equations involved, tilted photo refraction corrections can be conveniently inserted into the iterative loop.

Second, the foregoing discussion of atmospheric refraction in a tilted photograph assumes that the tilt angle is the angle between the optical axis and a vertical line. In analytical photogrammetry however, a local vertical coordinate system is generally used for object space (see Sec. 5-5). As a result, tilt angles will be related to the direction of the local vertical Z axis which is generally different from the true vertical direction. This effect however, is negligible for practical situations. Unless a photograph is more than about 300 km from the local vertical origin, the effect from ignoring this difference in vertical direction will generally be less than 0.001 mm.

References

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Bomford, G.: *Geodesy*, 4th ed., Clarendon Press, Oxford, 1980.

Hallert, B.: "Quality of Exterior Orientation," *Photogrammetric Engineering*, vol. 32, no. 3, 1966, p. 464.

Jones, A.D.: "The Development of the Wild Rectifiers," *Photogrammetric Record*, vol. 2, 1966, p. 181.

Keller, M., and G.C. Tewinkel: *Space Resection in Photogrammetry*, ESSA Technical Report C&GS 32, U.S. Coast and Geodetic Survey, Washington, 1966.

Mugnier, C.J.: "Analytical Rectification Using Artificial Points," *Photogrammetric Engineering and Remote Sensing*, vol. 44, no. 5, 1978, p. 579.

Trachsel, A.F.: "Electro-Optical Rectifier," *Photogrammetric Engineering*, vol. 33, no. 5, 1967, p. 513.
Wilson, K.R., and J. Vlcek: "Analytical Rectification," *Photogrammetric Engineering*, vol. 36, no. 6, 1970, p. 570.

Problems

10-1. A particular tilted aerial photograph exposed with a 152.047-mm-focal-length camera has a tilt angle of 3.25° and a swing angle of 135.00°. On this photograph, what are the auxiliary x' and y' photo coordinates for points a and b, whose photo coordinates measured with respect to the fiducial axes are $x_a = -62.41$ mm, $y_a = 76.80$ mm, $x_b = -98.75$ mm, and $y_b = -6.23$ mm?

10-2. Repeat Prob. 10-1, except that the camera focal length is 109.907 mm, tilt angle is 2.93°, swing angle is 235.50°, $x_a = 54.15$ mm, $y_a = 76.80$ mm, $x_b = 98.55$ mm, and $y_b = -12.06$ mm.

10-3. Calculate photographic scale for image points a and b of Prob. 10-1 if flying height above datum was 2039 m, and if elevations of points A and B were 306 m and 268 m above datum, respectively.

10-4. Calculate photographic scale for image points a and b of Prob. 10-2 if flying height above datum was 2239 m, and if elevations of points A and B were 329 m and 421 m above datum, respectively.

10-5. Illustrate and briefly describe the six elements of exterior orientation.

10-6. Name and briefly describe the three different methods of performing rectification. Discuss some advantages and disadvantages of each.

10-7. A tilted aerial photograph is exposed at $XL = 9274.2$ m, $YL = 8292.0$ m, and $ZL = 1500.1$ m. A rectified and ratioed photograph is to be produced with an optical-mechanical rectifier using a trial and error approach. The following table lists X, Y, and Z coordinates of five control points which will control the rectification. If the plane of rectification (average terrain) is at $h_{avg} = 110.0$ m, compute the displaced image locations X' and Y' of the control points as illustrated in Fig. 10-8.

Point	X, m	Y, m	Z, m
1	8,829.8	9,590.4	110.2
2	10,334.6	9,279.9	122.8
3	9,045.4	8,613.9	100.5
4	10,175.0	8,065.3	115.7
5	9473.7	7983.3	101.5

10-8. Repeat Prob. 10-7, except that the tilted photograph is exposed at $X_L = 12{,}732.6$ m, $Y_L = 15{,}820.0$ m, and $Z_L = 1{,}010.5$ m, average terrain is at $h_{avg} = 225.0$ m, and the X, Y, and Z coordinates of the control points are as listed in the following table.

Point	X, m	Y, m	Z, m
1	12,273.4	15,978.4	215.2
2	13,227.0	16,362.7	206.1
3	12,502.2	15,311.0	212.0
4	13,093.9	15,625.2	221.2
5	12,672.6	16,183.7	276.6

10-9. The following table lists measured photo coordinates for images of the control points 1-5 of Prob. 10-7 as well as additional points 6-8. Using the analytical rectification technique discussed in Secs. 10-11 and C-9, determine the rectified coordinates of points 6, 7, and 8. Use the original X and Y coordinates from the table from Prob. 10-7 as control, not the X' and Y' coordinates adjusted for relief displacement.

Point	X, mm	Y, mm
1	−63.31	70.27
2	69.48	30.78
3	−29.88	−21.77
4	74.31	−80.99
5	15.77	−82.89
6	−24.14	63.25
7	56.25	91.20
8	−12.99	−70.98

10-10. Repeat Prob. 10-9, except use X' and Y' coordinates, adjusted for relief displacement, from Prob. 10-7 as control for the rectification.

10-11. The following table lists measured photo coordinates for images of the control points 1 through 5 of Prob. 10-8 as well as additional points 6, 7, and 8. Using the analytical rectification technique discussed in Secs. 10-11 and C-9, determine the rectified coordinates of points 6 through 8. Use the original X and Y coordinates from the table from Prob. 10-8 as control, not the X' and Y' coordinates adjusted for relief displacement.

Point	X, mm	Y, mm
1	–129.84	24.21
2	97.59	37.49
3	–39.12	–43.79
4	99.80	–28.51
5	–38.96	37.11
6	50.93	–76.45
7	50.88	–25.41
8	–12.69	25.27

10-12. Repeat Prob. 10-11, except use X' and Y' coordinates, adjusted for relief displacement, from Prob. 10-8 as control for the rectification.

10-13. A tilted aerial photograph taken from a flying height of 3323 m above mean sea level contains the image *a* of object point *A* at fiducial coordinates x_a = 46.067 mm and y_a = –78.092 mm. The angles of tilt, swing, and azimuth for the photograph are 2.93°, –49.58°, and 144.39°, respectively. If the elevation of object point *A*, h_A = 295 m above mean sea level and the camera had a focal length, f = 152.013 mm, compute the x' and y' coordinates of the point, corrected for atmospheric refraction in the tilted photo.

10-14. Repeat Prob. 10-13 except that the flying height is 6032 m above mean sea level, x_a = 100.087 mm, y_a = 103.45 mm, tilt = 4.90°, swing = –147.28°, azimuth = 54.00°, h_A = 167 m, and f = 153.097 mm.

10-15. A nearly-horizontal terrestrial photo (θ= 0°) was exposed with a phototheodolite having a focal length of 55.00 mm. Find the horizontal angle *ALB* at the exposure station subtended by points *A* and *B* if corresponding images *a* and *b* have photo coordinates of x_a = 43.25 mm, y_a = –19.67 mm, x_b = –27.23 mm, and y_b = 24.73 mm.

10-16. For the data of Prob. 10-15, calculate the vertical angles from the exposure station to points *A* and *B*.

10-17. Repeat Prob. 10-15, except that the camera focal length is 150.00 mm, and the measured photo coordinates are x_a = –89.23 mm, y_a = 45.48 mm, x_b = 45.12 mm, and y_b = –27.73 mm.

10-18. Calculate the vertical angles for points *A* and *B* of Prob. 10-17.

10-19. An oblique terrestrial photo was exposed with the camera axis depressed at an angle of θ = 4.92°. The camera focal length was 192.30 mm. Calculate the horizontal and vertical angles between the rays from the camera station to object points *A* and *B* if their images have oblique photo coordinates of x_a = –75.82 mm, y_a = –14.28 mm, x_b = 81.58 mm, and y_b = –59.95 mm.

10-20. Repeat Prob. 10-19, except that the depression angle was 15.50°, the camera focal length was 90.01 mm, and the oblique photo coordinates were x_a = 2.70 mm, y_a = –38.23 mm, x_b = 67.23 mm, and

y_b = 23.09 mm.

---

[1] An equivalent vertical photograph is an imaginary truly vertical photo taken from the same exposure station as the tilted photo, with the same camera. The isocenter lies on the principal line of the tilted photo where the bisector of the tilt angle, constructed from the exposure station, intersects the tilted photo plane. A plane constructed through the isocenter, parallel to the datum principal plane, is the plane of an equivalent vertical photograph.

CHAPTER 11

Introduction to Analytical Photogrammetry

11-1 Introduction

Analytical photogrammetry is a term used to describe the rigorous mathematical calculation of coordinates of points in object space based upon camera parameters, measured photo coordinates, and ground control. Unlike the elementary methods presented in earlier chapters, this process rigorously accounts for any tilts that exist in the photos. Analytical photogrammetry generally involves the solution of large, complex systems of redundant equations by the method of least squares. The concepts of analytical photogrammetry existed for many years before the applications of them were considered practical due to the heavy computational effort that they entail. The evolution of computer technology and photogrammetric software, however, made analytical photogrammetry a commonplace technique. Analytical photogrammetry forms the basis of many modern hardware and software systems, including: stereoplotters (analytical and softcopy), digital terrain model generation, orthophoto production, digital photo rectification, and aerotriangulation.

This chapter presents an introduction to some fundamental topics and elementary applications in analytical photogrammetry. The coverage here is limited to computations involving single photos and stereo-pairs. Later chapters in the book cover more advanced topics and applications in this subject. In particular, Chap. 17 describes analytical photogrammetry solutions for handling strips and blocks of photos.

11-2 Image Measurements

A fundamental type of measurement used in analytical photogrammetry is an x and y photo coordinate pair. These coordinates *must be related to the principal point as the origin*. Since mathematical relationships in analytical photogrammetry are based on assumptions such as "light rays travel in straight lines" and "the focal plane of a frame camera is flat," various coordinate refinements may be required to correct measured photo coordinates for distortion effects that otherwise cause these assumptions to be violated. Techniques of photo coordinate refinement, as presented in Secs. 4-7 through 4-12 and Sec. 10-16, can be used to correct photo coordinates so that the assumptions will be valid.

A number of instruments and techniques are available for making photo coordinate measurements, as discussed in Secs. 4-4 through 4-6. To ensure results having the highest accuracy, these measurements must be made with great care. In many analytical photogrammetry methods, it is necessary to measure image coordinates of common object points that appear in more than one photograph. In these cases it is essential that the image of each object point be precisely identified between photos so that the measurements will be consistent.

11-3 Control Points

In addition to measurement of image coordinates, a certain number of control points in object space are generally required for analytical photogrammetry. Object space coordinates of these control

points, which may be either image-identifiable features or exposure stations of the photographs themselves, are generally determined via some type of field survey technique such as GPS. It is important that the object space coordinates be based on a three-dimensional cartesian system which has straight, mutually perpendicular axes. This often requires three-dimensional coordinate conversions which are described in Chap. 5 and App. F.

11-4 Collinearity Condition

Perhaps the most fundamental and useful relationship in analytical photogrammetry is the *collinearity* condition. Collinearity, as described in App. D, is the condition that the exposure station, any object point, and its photo image all lie along a straight line in three-dimensional space. The collinearity condition is illustrated in Fig. 11-1, where L, a, and A lie along a straight line. Two equations express the collinearity condition for any point on a photo: one equation for the x photo coordinate and another for the y photo coordinate. The mathematical relationships are developed in App. D and expressed by Eqs. (D-5) and (D-6). They are repeated here for convenience.

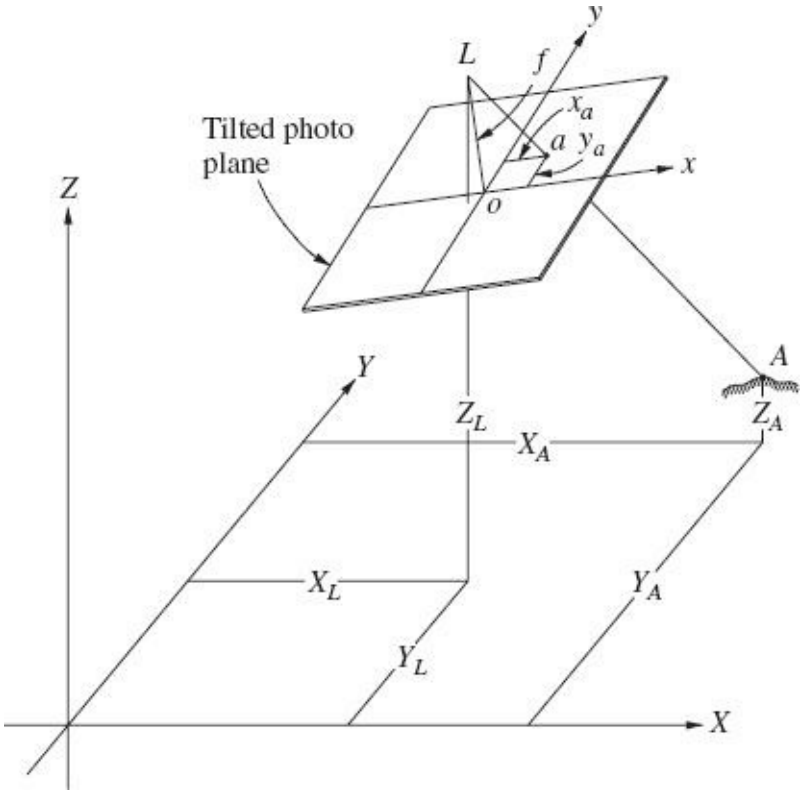


FIGURE 11-1 The *collinearity* condition.

$$x_a = x_o - f\left[\frac{m_{11}(X_A - X_L) + m_{12}(Y_A - Y_L) + m_{13}(Z_A - Z_L)}{m_{31}(X_A - X_L) + m_{32}(Y_A - Y_L) + m_{33}(Z_A - Z_L)}\right] \tag{11-1}$$

$$y_a = y_o - f\left[\frac{m_{21}(X_A - X_L) + m_{22}(Y_A - Y_L) + m_{23}(Z_A - Z_L)}{m_{31}(X_A - X_L) + m_{32}(Y_A - Y_L) + m_{33}(Z_A - Z_L)}\right] \tag{11-2}$$

In Eqs. (11-1) and (11-2), x_a and y_a are the photo coordinates of image point a; X_A, Y_A, and Z_A are object

space coordinates of point A; X_L, Y_L, and Z_L are object space coordinates of the exposure station; f is the camera focal length; $x_\circ$ and $y_\circ$ are the coordinates of the principal point (usually known from camera calibration); and the m's (as described in Sec. C-7) are functions of three rotation angles, and most often omega, phi, and kappa are the angles employed.

The collinearity equations are nonlinear and can be linearized by using Taylor's theorem as described in Sec. D-5. The linearized forms are Eqs. (D-15) and (D-16), and they are also repeated here for convenience.

$$b_{11}\,d\omega + b_{12}\,d\phi + b_{13}\,d\kappa - b_{14}\,dX_L - b_{15}\,dY_L - b_{16}\,dZ_L + b_{14}\,dX_A + b_{15}\,dY_A + b_{16}\,dZ_A = J + v_{x_a} \tag{11-3}$$

$$b_{21}\,d\omega + b_{22}\,d\phi + b_{23}\,d\kappa - b_{24}\,dX_L - b_{25}\,dY_L - b_{26}\,dZ_L + b_{24}\,dX_A + b_{25}\,dY_A + b_{26}\,dZ_A = K + v_{y_a} \tag{11-4}$$

In Eqs. (11-3) and (11-4), v_{x_a} and v_{y_a} are residual errors in measured x_a and y_a image coordinates; $d\omega$, $d\Phi$, and $d\kappa$ are corrections to initial approximations for the orientation angles of the photo; dX_L, dY_L, and dZ_L are corrections to initial approximations for the exposure station coordinates; and dX_A, dY_A, and dZ_A are corrections to initial values for the object space coordinates of point A. The b's and the J and K terms are described in Sec. D-5. Because higher order terms are ignored in linearization by Taylor's theorem, the linearized forms of the equations are approximations. They must therefore be solved iteratively, as described in App. D, until the magnitudes of corrections to initial approximations become negligible.

11-5 Coplanarity Condition

Coplanarity, as illustrated in Fig. 11-2, is the condition that the two exposure stations of a stereopair, any object point, and its corresponding image points on the two photos all lie in a common plane. In the figure, for example, points L_1, L_2, a_1, a_2, and A all lie in the same plane. The coplanarity condition is developed in Sec. D-7 and shown in Eq. (11-5).

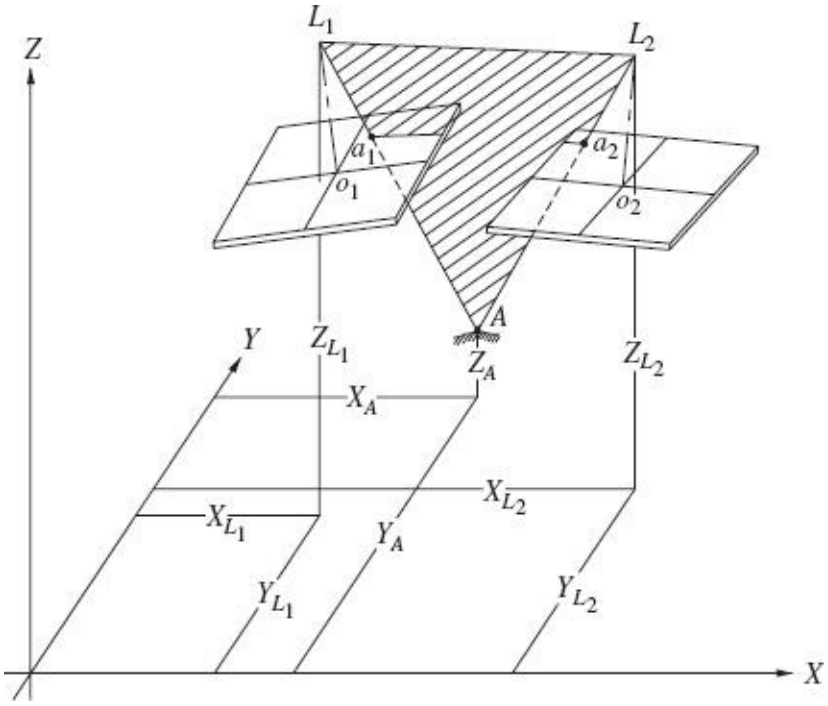


FIGURE 11-2 The *coplanarity* condition.

$$0 = B_X(E_1F_2 - E_2F_1) + B_Y(F_1D_2 - F_2D_1) + B_Z(D_1E_2 - D_2E_1) \tag{11-5}$$

In Eq. (11-5), subscripts 1 and 2 affixed to terms *D*, *E*, and *F* indicate that the terms apply to either photo 1 or photo 2. The *m*'s again are functions of the three rotation angles omega, phi, and kappa, as defined in Sec. C-7. One coplanarity equation may be written for each object point whose images appear on both photos of the stereopair. The coplanarity equations do not contain object space coordinates as unknowns; rather, they contain only the elements of exterior orientation of the two photos of the stereopair.

Like collinearity equations, the coplanarity equation is nonlinear and must be linearized by using Taylor's theorem and solved iteratively for corrections to approximations of the orientation parameters. Linearization of the coplanarity equation is described in Sec. D-8.

11-6 Space Resection by Collinearity

Space resection, as discussed in Sec. 10-10, is a method of determining the six elements of exterior orientation ($\omega, \phi, \kappa, X_L, Y_L$, and Z_L) of a photograph. This method requires a minimum of three control points, with known *XYZ* object space coordinates, to be imaged in the photograph. If the ground control coordinates are assumed to be known and fixed, then the linearized forms of the space resection collinearity equations for a point *A* are

$$b_{11}\,d\omega + b_{12}\,d\phi + b_{13}\,d\kappa - b_{14}\,dX_L - b_{15}\,dY_L - b_{16}\,dZ_L = J + v_{x_a} \tag{11-6}$$

$$b_{21}\,d\omega + b_{22}\,d\phi + b_{23}\,d\kappa - b_{24}\,dX_L - b_{25}\,dY_L - b_{26}\,dZ_L = K + v_{y_a} \tag{11-7}$$

In Eqs. (11-6) and (11-7), the terms are as defined in Secs. 11-4 and D-5. Two equations are formed

for each control point, which gives six equations if the minimum of three control points is used. In this case a unique solution results for the six unknowns, and the residual terms on the right sides of Eqs. (11-6) and (11-7) will be zero. If four or more control points are used, more than six equations can be formed, allowing a least squares solution.

Since the collinearity equations are nonlinear, and have been linearized using Taylor's theorem, initial approximations are required for the unknown orientation parameters. For the typical case of near-vertical photography, zero values can be used as initial approximations for ω and ϕ. The value of Z_L (the height H above datum) can be computed using the method discussed in Sec. 6-9. Since this method requires only two control points, several solutions are possible, using different pairs of control points. An improved approximation can be made by computing several values for H and taking the average. After H has been determined, ground coordinates from a vertical photograph, as described in Sec. 6-7, can be computed, using the measured x and y photo coordinates, focal length f, flying height H, and elevation of the object point Z [or h as it appears in Eqs. (6-5) and (6-6)]. A two-dimensional conformal coordinate transformation can then be performed, which relates the ground coordinates as computed from the vertical photo equations to the control values. The two-dimensional conformal coordinate transformation Eqs. (C-12), repeated here in a slightly different form, are used.

$$X = ax' - by' + T_X$$

$$Y = ay' + bx' + T_Y \tag{11-8}$$

In Eqs. (11-8), X and Y are ground control coordinates for the point; x' and y' are ground coordinates from a vertical photograph as computed by Eqs. (6-5) and (6-6); and a, b, T_X, and T_Y are the transformation parameters. A pair of equations of the type of Eqs. (11-8) can be written for each control point, and the four unknown parameters computed by least squares (see Secs. C-4 and C-5). The translation factors T_X and T_Y determined from this solution can then be used as initial approximations for X_L and Y_L, respectively. Rotation angle q, which can be computed by Eq. (C-11), can be used as an approximation for k.

By using these initial approximations in Eqs. (11-6) and (11-7), a least squares solution can be computed for the unknown corrections to the approximations. The solution is iterated until the corrections become negligible.

Example 11-1

A near-vertical aerial photograph taken with a 152.916-mm-focal-length camera contains images of four ground control points A through D. Refined photo coordinates and ground control coordinates (in a local vertical system) of the four points are listed in the following table. Calculate the exterior orientation parameters $\boldsymbol{\omega}$, $\boldsymbol{\phi}$, $\boldsymbol{\kappa}$, X_L, Y_L, and Z_L for this photograph.

	Photo Coordinates		Ground Control Coordinates		
Point	x, mm	y, mm	X, m	Y, m	Z, m
A	86.421	−83.977	1,268.102	1,455.027	22.606
B	−100.916	92.582	732.181	545.344	22.299
C	−98.322	−89.161	1,454.553	731.666	22.649
D	78.812	98.123	545.245	1,268.232	22.336

Solution Since the computations involved in analytical photogrammetry—even in the most elementary problems—are too difficult and time consuming for hand solution, one should use a computer program to solve them. The following describes the steps and shows the output from a space resection program.

1. With an ASCII Editor, create the following file with a ".dat" extension.

```
152.916
A   86.421 -83.977  1,268.102  1,455.027  22.606
B -100.916  92.582    732.181    545.344  22.299
C  -98.322 -89.161  1,454.553    731.666  22.649
D   78.812  98.123    545.245  1,268.232  22.336
```

   The first line is the focal length of the camera. The next four lines consist of the point name, the *x* and *y* photo coordinates, and the ground control coordinates for each point.
2. Running the resect program yields an output file containing the following:

```
Exterior orientation parameters:

Omega =    -0.4109 +- 0.0009 deg
Phi   =     1.2101 +- 0.0009 deg
Kappa =   102.8003 +- 0.0004 deg
XL    =  1027.8571 +- 0.0134
YL    =  1044.1138 +- 0.0134
ZL    =   648.1974 +- 0.0042

Photo coordinate residuals:

   point   x-res   y-res
       A  0.0006 -0.0023
       B -0.0001  0.0024
       C  0.0021  0.0003
       D -0.0026 -0.0004

 average -0.0000  0.0000
     RMS  0.0017  0.0017

Standard error of unit weight: 0.0017
Degrees of freedom: 2
```

   The solved exterior orientation parameters for the angular and positional components with standard deviations are shown in degrees and meters, respectively. The photo coordinate residuals are in millimeters. ▲

11-7 Space Intersection by Collinearity

If space resection is used to determine the elements of exterior orientation for both photos of a stereopair, as described in the preceding section, then object point coordinates for points that lie in the stereo overlap area can be calculated. The procedure is known as *space intersection*, so called because corresponding rays to the same object point from the two photos must intersect at the point, as shown in Fig. 11-3. To calculate the coordinates of point *A* by space intersection, collinearity equations of the linearized form given by Eqs. (11-3) and (11-4) can be written for each new point, such as point *A* of Fig. 11-3. Note, however, that since the six elements of exterior orientation are known, the only remaining unknowns in these equations are dX_A, dY_A, and dZ_A. These are corrections to be applied to initial approximations for object space coordinates X_A, Y_A, and Z_A, respectively, for ground point *A*.

The linearized forms of the space intersection equations for point *A* are

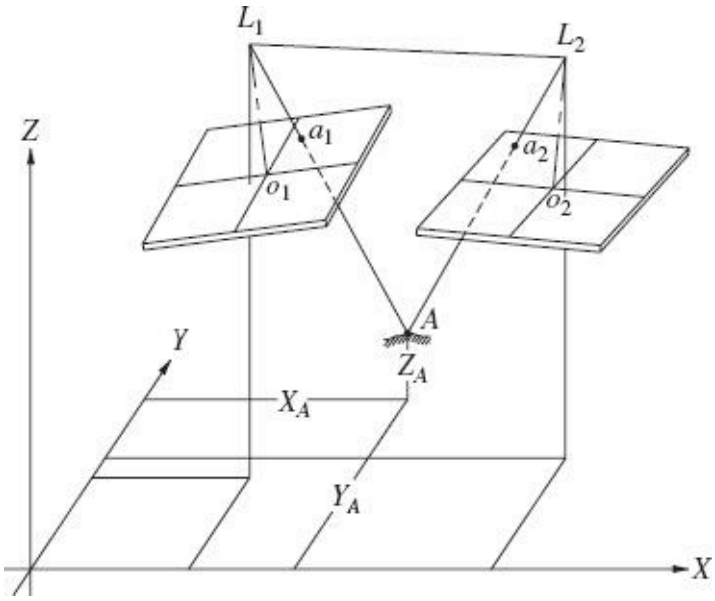


FIGURE 11-3 Space intersection with a stereopair of aerial photos.

$$b_{14}\,dX_A + b_{15}\,dY_A + b_{16}\,dZ_A = J + v_{x_a} \tag{11-9}$$

$$b_{24}\,dX_A + b_{25}\,dY_A + b_{26}\,dZ_A = K + v_{y_a} \tag{11-10}$$

In Eqs. (11-9) and (11-10), the terms are as defined in Secs. 11-4 and D-5. Two equations of this form can be written for point a_1 of the left photo, and two more for point a_2 of the right photo; hence four equations result, and the three unknowns dX_A, dY_A, and dZ_A can be computed in a least squares solution. These corrections are added to the initial approximations to obtain revised values for X_A, Y_A, and Z_A. The solution is then repeated until the magnitudes of the corrections become negligible.

Again because the equations have been linearized using Taylor's theorem, initial approximations are required for each point whose object space coordinates are to be computed. For these calculations, with normal aerial photography vertical photos can be assumed, and the initial approximations can be determined by using the parallax equations [Eqs. (8-5) through (8-7)]. Note that because the *X*, *Y*, and *Z* coordinates for both exposure stations are known, for making these computations *H* can be taken as the average of Z_{L_1} and Z_{L_2}, and *B* is computed from

$$B = \sqrt{(X_{L_2} - X_{L_1})^2 + (Y_{L_2} - Y_{L_1})^2} \tag{a}$$

The coordinates that result from Eqs. (8-6) and (8-7) are in the arbitrary system described in Sec. 8-6. (Let these coordinates be designated as x' and y'.) To convert them to the *X* and *Y* ground system, coordinate transformation Eqs. (11-8) can be used. For this transformation, the two exposure stations can serve as the control because their *X* and *Y* coordinates are known in the ground system, and their x' and y' coordinates in the parallax system are $x'_{L_1} = y'_{L_1} = y'_{L_2} = 0$, and $x'_{L_2} = B$. Since there are four equations and four unknowns, the transformation parameters can be solved for directly and applied to the

imaged points to get the horizontal coordinate initial approximations as shown in Eqs. (11-11).

$$X_A = \frac{X_{L_2} - X_{L_1}}{B} x'_a - \frac{Y_{L_2} - Y_{L_1}}{B} y'_a + X_{L_1}$$

$$Y_A = \frac{X_{L_2} - X_{L_1}}{B} y'_a + \frac{Y_{L_2} - Y_{L_1}}{B} x'_a + Y_{L_1} \qquad (11\text{-}11)$$

Example 11-2

A stereopair of images taken with a 152.057-mm-focal-length camera has exterior orientation parameters shown in the first table below. The images of four points have measured photo coordinates shown in the second table below. Use the analytical space intersection program to find the ground coordinates of each of the points.

Image	ω,°	ϕ,°	κ,°	X_L, m	Y_L, m	Z_L, m
1	1.4022	−0.3112	0.6470	9,577.252	10,214.285	555.192
2	−0.1557	−1.7063	0.5513	9,803.241	10,219.622	556.601

	Image 1 Photo Coordinates		Image 2 Photo Coordinates	
Point	x, mm	y, mm	x, mm	y, mm
A	−12.843	−54.155	−87.550	−51.157
B	−22.720	−4.828	−98.272	−1.702
C	−0.433	70.321	−75.465	75.310
D	−25.993	−58.086	−101.077	−55.213

Solution

1. With an ASCII editor, create a file with a ".dat" extension in the following way:

```
152.057
 1.4022 -0.3112  0.6470 9,577.252 10214.285 555.192
-0.1557 -1.7063  0.5513 9,803.241 10219.622 556.601
a -12.843 -54.155    -87.550 -51.157
b -22.720  -4.828    -98.272  -1.702
c -0.433  70.321    -75.465  75.310
d -25.993 -58.086   -101.077 -55.213
```

   The first line is the focal length, the second two lines contain the exterior orientation parameters of the photos, and the remaining lines contain the point name and x and y coordinates of the point on each photo.
2. Run the intersect program, producing the following output file:

Object Space Coordinates:

Name	X	Y	Z	SD-X	SD-Y	SD-Z
a	9,540.831	10,052.850	65.294	0.037	0.039	0.096
b	9,507.255	10,209.928	67.164	0.002	0.001	0.005
c	9,575.948	10,454.966	66.438	0.003	0.004	0.008
d	9,499.192	10,040.285	66.068	0.009	0.008	0.020

Photo Coordinate Residuals

Name	xl-res	yl-res	xr-res	yr-res
a	-0.000	+0.007	+0.000	-0.007
b	+0.000	-0.000	-0.000	+0.000
c	+0.000	-0.001	-0.000	+0.001
d	-0.000	+0.001	+0.000	-0.001

The object space, or ground coordinates and their corresponding standard deviations are in meters. The photo coordinate residuals are in millimeters. The *x* photo coordinate residuals are nearly equal to zero for both the left and right photos. This is expected for a stereopair displaced in the *x* direction because *x* parallax determines the *Z* coordinate in object space. The *y* photo coordinate residuals indicate the presence of *y* parallax, and their sizes are good indicators of the quality of the solution. Generally, *y* coordinate residuals less than about 5 or 6 micrometers (μm) indicate a rather good solution. ▲

11-8 Analytical Stereomodel

Aerial photographs for most applications are taken so that adjacent photos overlap by more than 50 percent. Two adjacent photographs that overlap in this manner form a *stereopair*, and object points that appear in the overlap area constitute a *stereomodel*. The mathematical calculation of three-dimensional ground coordinates of points in the stereomodel by analytical photogrammetric techniques forms an *analytical stereomodel*.

The process of forming an analytical stereomodel involves three primary steps: *interior orientation, relative orientation*, and *absolute orientation*. After these three steps are achieved, points in the analytical stereomodel will have object coordinates in the ground coordinate system. These points can then be used for many purposes, such as digital mapping, serving as control for orthophoto production, or DEM generation. The three orientation steps can be performed as distinct mathematical operations, or it is possible to combine them in a simultaneous solution. In this chapter, the three steps will be described as distinct, sequential operations.

11-9 Analytical Interior Orientation

Interior orientation for analytical photogrammetry is the step which mathematically recreates the geometry that existed in the camera when a particular photograph was exposed. This requires camera calibration information as well as quantification of the effects of atmospheric refraction. These procedures, commonly called *photo coordinate refinement*, are described in Secs. 4-9 through 4-12. In the case of film photography, the process begins with coordinates of fiducials and image points which have been measured by a comparator or related device. A two-dimensional (usually affine) coordinate transformation can be used to relate the comparator, or softcopy coordinates to the fiducial coordinate system as well as to correct for film distortion. There is no need for fiducial measurements when using photographs from a digital camera, since it has a consistent coordinate system in the form of pixels. For both film and digital photography the lens distortion and principal-point information from camera calibration are then used to refine the coordinates so that they are correctly related to the

principal point and free from lens distortion. Finally, atmospheric refraction corrections can be applied to the photo coordinates to complete the refinement and, therefore, finish the interior orientation.

11-10 Analytical Relative Orientation

Analytical relative orientation is the process of determining the relative angular attitude and positional displacement between the photographs that existed when the photos were taken. This involves defining certain elements of exterior orientation and calculating the remaining ones. The resulting exterior orientation parameters will not be the actual values that existed when the photographs were exposed; however, they will be correct in a "relative sense" between the photos. In analytical relative orientation, it is common practice to fix the exterior orientation elements ω, ϕ, κ, X_L, and Y_L of the left photo of the stereopair to zero values. Also for convenience, Z_L of the left photo (Z_{L_1}) is set equal to f, and X_L of the right photo (X_{L_2}) is set equal to the photo base b. (With these choices for Z_{L_1} and X_{L_2}, initial approximations for the unknowns are more easily calculated, as will be explained later.) This leaves five elements of the right photo that must be determined. Figure 11-4 illustrates a stereomodel formed by analytical relative orientation.

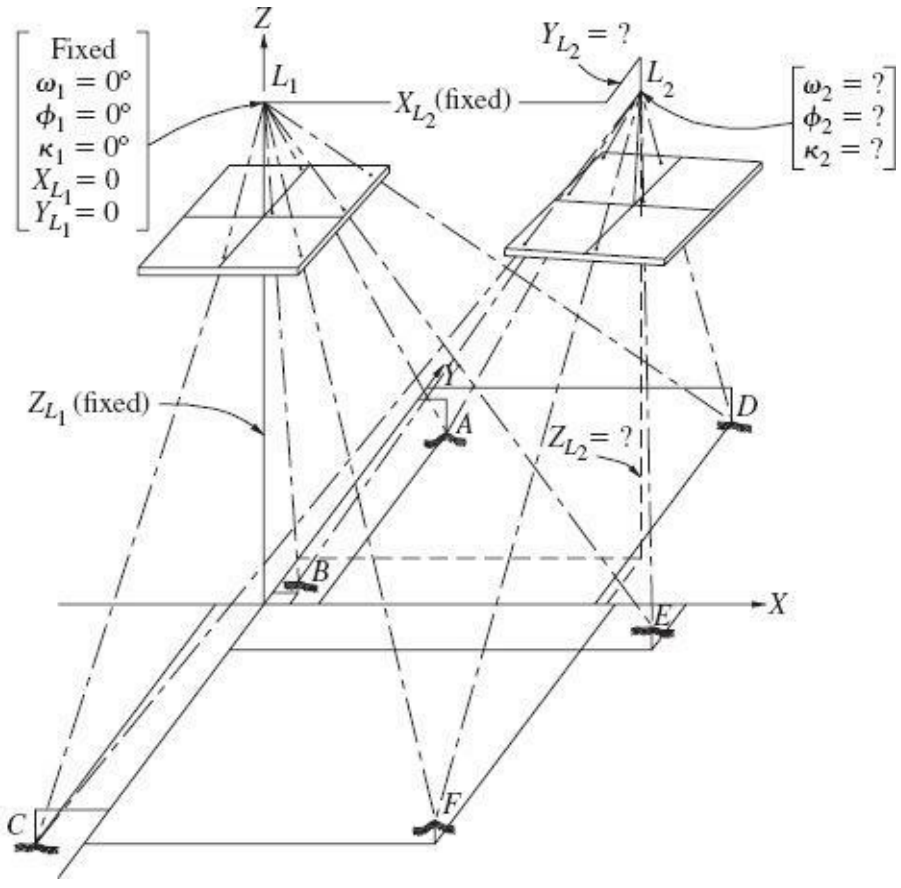


FIGURE 11-4 Analytical relative orientation of a stereopair.

Although the coplanarity condition equation can be used for analytical relative orientation, the collinearity condition is more commonly applied. In applying collinearity, each object point in the stereomodel contributes four equations: a pair of equations for the x and y photo coordinates of its image in the left photo, and a pair of equations for the x and y photo coordinates of its image in the

right photo. In addition to the five unknown orientation elements, each object point adds three more unknowns which are their *X*, *Y*, and *Z* coordinates in the stereomodel. Thus each point used in relative orientation results in a net gain of one equation for the overall solution, and therefore at least five object points are required for a solution. If six or more points are available, an improved solution is possible through least squares. If six points were used for relative orientation, a system of 24 equations with 23 unknowns would result; and if 12 points were used, the system would consist of 48 equations and 41 unknowns.

$$
A = \begin{array}{c}
\begin{array}{cccccccccccc}
d\omega_2 & d\phi_2 & d\kappa_2 & dY_{L_2} & dZ_{L_2} & dX_A & dY_A & dZ_A & dX_B & dY_B & dZ_B
\end{array} \\
\left[\begin{array}{cccccccccccc}
0 & 0 & 0 & 0 & 0 & (b_{a_{14}})_1 & (b_{a_{15}})_1 & (b_{a_{16}})_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & (b_{a_{24}})_1 & (b_{a_{25}})_1 & (b_{a_{26}})_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (b_{b_{14}})_1 & (b_{b_{15}})_1 & (b_{b_{16}})_1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (b_{b_{24}})_1 & (b_{b_{25}})_1 & (b_{b_{26}})_1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{a_{11}})_2 & (b_{a_{12}})_2 & (b_{a_{13}})_2 & (-b_{a_{15}})_2 & (-b_{a_{16}})_2 & (b_{a_{14}})_2 & (b_{a_{15}})_2 & (b_{a_{16}})_2 & 0 & 0 & 0 \\
(b_{a_{21}})_2 & (b_{a_{22}})_2 & (b_{a_{23}})_2 & (-b_{a_{25}})_2 & (-b_{a_{26}})_2 & (b_{a_{24}})_2 & (b_{a_{25}})_2 & (b_{a_{26}})_2 & 0 & 0 & 0 \\
(b_{b_{11}})_2 & (b_{b_{12}})_2 & (b_{b_{13}})_2 & (-b_{b_{15}})_2 & (-b_{b_{16}})_2 & 0 & 0 & 0 & (b_{b_{14}})_2 & (b_{b_{15}})_2 & (b_{b_{16}})_2 \\
(b_{b_{21}})_2 & (b_{b_{22}})_2 & (b_{b_{23}})_2 & (-b_{b_{25}})_2 & (-b_{b_{26}})_2 & 0 & 0 & 0 & (b_{b_{24}})_2 & (b_{b_{25}})_2 & (b_{b_{26}})_2 \\
(b_{c_{11}})_2 & (b_{c_{12}})_2 & (b_{c_{13}})_2 & (-b_{c_{15}})_2 & (-b_{c_{16}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{c_{21}})_2 & (b_{c_{22}})_2 & (b_{c_{23}})_2 & (-b_{c_{25}})_2 & (-b_{c_{26}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{d_{11}})_2 & (b_{d_{12}})_2 & (b_{d_{13}})_2 & (-b_{d_{15}})_2 & (-b_{d_{16}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{d_{21}})_2 & (b_{d_{22}})_2 & (b_{d_{23}})_2 & (-b_{d_{25}})_2 & (-b_{d_{26}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{e_{11}})_2 & (b_{e_{12}})_2 & (b_{e_{13}})_2 & (-b_{e_{15}})_2 & (-b_{e_{16}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{e_{21}})_2 & (b_{e_{22}})_2 & (b_{e_{23}})_2 & (-b_{e_{25}})_2 & (-b_{e_{26}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{f_{11}})_2 & (b_{f_{12}})_2 & (b_{f_{13}})_2 & (-b_{f_{15}})_2 & (-b_{f_{16}})_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
(b_{f_{21}})_2 & (b_{f_{22}})_2 & (b_{f_{23}})_2 & (-b_{f_{25}})_2 & (-b_{f_{26}})_2 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
\end{array}
$$

Prior to solving the linearized collinearity equations, initial approximations for all unknown values must be determined. For photography that was intended to be vertical, values of zero are commonly used for initial estimates of ω_2, $\boldsymbol{\phi_2}$, $\boldsymbol{\kappa_2}$, and Y_{L2}. An initial value for Z_{L_2} may be selected equal to the value used for Z_{L_1}. If the constraints that were noted earlier are used for the parameters, that is, $\boldsymbol{\omega_1} = \boldsymbol{\phi_1} = \boldsymbol{\kappa_1} = X_{L_1} = Y_{L_1} = 0$, $Z_{L_1} = f$, and $X_{L_2} = b$, then the scale of the stereomodel is approximately equal to photo scale. Thus the x and y photo coordinates of the left photo are good approximations for X and Y object space coordinates, and zeros are good approximations for Z object space coordinates, respectively.

Suppose that the six points of Fig. 11-4 were used in analytical relative orientation. In matrix form, the system of 24 equations involving 23 unknowns could be expressed as follows:

$$ {}_{24}A^{23}{}_{23}X^{1} = {}_{24}L^{1} + {}_{24}V^{1} \tag{11-12} $$

where

dX_C	dY_C	dZ_C	dX_D	dY_D	dZ_D	dX_E	dY_E	dZ_E	dX_F	dY_F	dZ_F
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
$(b_{c_{14}})_1$	$(b_{c_{15}})_1$	$(b_{c_{16}})_1$	0	0	0	0	0	0	0	0	0
$(b_{c_{24}})_1$	$(b_{c_{25}})_1$	$(b_{c_{26}})_1$	0	0	0	0	0	0	0	0	0
0	0	0	$(b_{d_{14}})_1$	$(b_{d_{15}})_1$	$(b_{d_{16}})_1$	0	0	0	0	0	0
0	0	0	$(b_{d_{24}})_1$	$(b_{d_{25}})_1$	$(b_{d_{26}})_1$	0	0	0	0	0	0
0	0	0	0	0	0	$(b_{e_{14}})_1$	$(b_{e_{15}})_1$	$(b_{e_{16}})_1$	0	0	0
0	0	0	0	0	0	$(b_{e_{24}})_1$	$(b_{e_{25}})_1$	$(b_{e_{26}})_1$	0	0	0
0	0	0	0	0	0	0	0	0	$(b_{f_{14}})_1$	$(b_{f_{15}})_1$	$(b_{f_{16}})_1$
0	0	0	0	0	0	0	0	0	$(b_{f_{24}})_1$	$(b_{f_{25}})_1$	$(b_{f_{26}})_1$
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
$(b_{c_{14}})_2$	$(b_{c_{15}})_2$	$(b_{c_{16}})_2$	0	0	0	0	0	0	0	0	0
$(b_{c_{24}})_2$	$(b_{c_{25}})_2$	$(b_{c_{26}})_2$	0	0	0	0	0	0	0	0	0
0	0	0	$(b_{d_{14}})_2$	$(b_{d_{15}})_2$	$(b_{d_{16}})_2$	0	0	0	0	0	0
0	0	0	$(b_{d_{24}})_2$	$(b_{d_{25}})_2$	$(b_{d_{26}})_2$	0	0	0	0	0	0
0	0	0	0	0	0	$(b_{e_{14}})_2$	$(b_{e_{15}})_2$	$(b_{e_{16}})_2$	0	0	0
0	0	0	0	0	0	$(b_{e_{24}})_2$	$(b_{e_{25}})_2$	$(b_{e_{26}})_2$	0	0	0
0	0	0	0	0	0	0	0	0	$(b_{f_{14}})_2$	$(b_{f_{15}})_2$	$(b_{f_{16}})_2$
0	0	0	0	0	0	0	0	0	$(b_{f_{24}})_2$	$(b_{f_{25}})_2$	$(b_{f_{26}})_2$

$$X = \begin{bmatrix} d\omega_2 \\ d\phi_2 \\ d\kappa_2 \\ dY_{L_2} \\ dZ_{L_2} \\ dX_A \\ dY_A \\ dZ_A \\ dX_B \\ dY_B \\ dZ_B \\ dX_C \\ dY_C \\ dZ_C \\ dX_D \\ dY_D \\ dZ_D \\ dX_E \\ dY_E \\ dZ_E \\ dX_F \\ dY_F \\ dZ_F \end{bmatrix} \quad L = \begin{bmatrix} (J_a)_1 \\ (K_a)_1 \\ (J_b)_1 \\ (K_b)_1 \\ (J_c)_1 \\ (K_c)_1 \\ (J_d)_1 \\ (K_d)_1 \\ (J_e)_1 \\ (K_e)_1 \\ (J_f)_1 \\ (K_f)_1 \\ (J_a)_2 \\ (K_a)_2 \\ (J_b)_2 \\ (K_b)_2 \\ (J_c)_2 \\ (K_c)_2 \\ (J_d)_2 \\ (K_d)_2 \\ (J_e)_2 \\ (K_e)_2 \\ (J_f)_2 \\ (K_f)_2 \end{bmatrix} \quad V = \begin{bmatrix} (v_{x_a})_1 \\ (v_{y_a})_1 \\ (v_{x_b})_1 \\ (v_{y_b})_1 \\ (v_{x_c})_1 \\ (v_{y_c})_1 \\ (v_{x_d})_1 \\ (v_{y_d})_1 \\ (v_{x_e})_1 \\ (v_{y_e})_1 \\ (v_{x_f})_1 \\ (v_{y_f})_1 \\ (v_{x_a})_2 \\ (v_{y_a})_2 \\ (v_{x_b})_2 \\ (v_{y_b})_2 \\ (v_{x_c})_2 \\ (v_{y_c})_2 \\ (v_{x_d})_2 \\ (v_{y_d})_2 \\ (v_{x_e})_2 \\ (v_{y_e})_2 \\ (v_{x_f})_2 \\ (v_{y_f})_2 \end{bmatrix}$$

Note that Eq. (11-12) conforms to Eq. (B-10).

The terms of these matrices are from Eqs. (11-3) and (11-4), and the method of calculating each is explained in Sec. D-5. The subscripts *a*, *b*, *c*, *d*, *e*, and *f* correspond to the point names; subscript 1 refers to the left photo; and subscript 2 refers to the right photo. Note that the elements of the *X* matrix are listed at the tops of the columns of the *A* matrix for illustration purposes only.

Upon studying these matrices, particularly the *A* matrix, their systematic nature becomes apparent. The fact that many submatrices of zeros exist also indicates the relative ease with which partitioning could be applied in the solution. The following equations show the form of the *A*, *X*, *L*,

and V matrices when partitioned by the standard approach.

$$A = \begin{bmatrix}
{}_2 0^5 & {}_2\ddot{B}^3_{1a} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 \\
{}_2 0^5 & {}_2 0^3 & {}_2\ddot{B}^3_{1b} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 \\
{}_2 0^5 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{1c} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 \\
{}_2 0^5 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{1d} & {}_2 0^3 & {}_2 0^3 \\
{}_2 0^5 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{1e} & {}_2 0^3 \\
{}_2 0^5 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{1f} \\
{}_2\dot{B}^5_{2a} & {}_2\ddot{B}^3_{2a} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 \\
{}_2\dot{B}^5_{2b} & {}_2 0^3 & {}_2\ddot{B}^3_{2b} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 \\
{}_2\dot{B}^5_{2c} & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{2c} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 \\
{}_2\dot{B}^5_{2d} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{2d} & {}_2 0^3 & {}_2 0^3 \\
{}_2\dot{B}^5_{2e} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{2e} & {}_2 0^3 \\
{}_2\dot{B}^5_{2f} & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2 0^3 & {}_2\ddot{B}^3_{2f}
\end{bmatrix} \quad X = \begin{bmatrix} \dot{\Delta}_2 \\ \ddot{\Delta}_A \\ \ddot{\Delta}_B \\ \ddot{\Delta}_C \\ \ddot{\Delta}_D \\ \ddot{\Delta}_E \\ \ddot{\Delta}_F \end{bmatrix} \quad L = \begin{bmatrix} \varepsilon_{1a} \\ \varepsilon_{1b} \\ \varepsilon_{1c} \\ \varepsilon_{1d} \\ \varepsilon_{1e} \\ \varepsilon_{1f} \\ \varepsilon_{2a} \\ \varepsilon_{2b} \\ \varepsilon_{2c} \\ \varepsilon_{2d} \\ \varepsilon_{2e} \\ \varepsilon_{2f} \end{bmatrix} \quad V = \begin{bmatrix} v_{1a} \\ v_{1b} \\ v_{1c} \\ v_{1d} \\ v_{1e} \\ v_{1f} \\ v_{2a} \\ v_{2b} \\ v_{2c} \\ v_{2d} \\ v_{2e} \\ v_{2f} \end{bmatrix}$$

where

$${}_2 0^5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad {}_2 0^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$${}_2\dot{B}^5_{2p} = \begin{bmatrix} (b_{p_{11}})_2 & (b_{p_{12}})_2 & (b_{p_{13}})_2 & (-b_{p_{15}})_2 & (-b_{p_{16}})_2 \\ (b_{p_{21}})_2 & (b_{p_{22}})_2 & (b_{p_{23}})_2 & (-b_{p_{25}})_2 & (-b_{p_{26}})_2 \end{bmatrix} \quad {}_2\ddot{B}^3_{ip} = \begin{bmatrix} (b_{p_{14}})_i & (b_{p_{15}})_i & (b_{p_{16}})_i \\ (b_{p_{24}})_i & (b_{p_{25}})_i & (b_{p_{26}})_i \end{bmatrix}$$

$$\dot{\Delta}_2 = \begin{bmatrix} d\omega_2 \\ d\phi_2 \\ d\kappa_2 \\ dY_{L_2} \\ dZ_{L_2} \end{bmatrix} \quad \ddot{\Delta}_p = \begin{bmatrix} dX_p \\ dY_p \\ dZ_p \end{bmatrix} \quad \varepsilon_{ip} = \begin{bmatrix} (J_p)_i \\ (K_p)_i \end{bmatrix} \quad v_{ip} = \begin{bmatrix} (v_{x_p})_i \\ (v_{y_p})_i \end{bmatrix}$$

In the above submatrices, p is the point designation, and i is the photo designation. The prefixed subscript and postfixed superscript designate the number of rows and columns, respectively.

When the observation equations are partitioned in the above-described manner, the least squares solution $(A^{\mathrm{T}}A)X = (A^{\mathrm{T}}L)$ takes the following form.

$$\begin{bmatrix} \dot{N}_2 & \bar{N}_a & \bar{N}_b & \bar{N}_c & \bar{N}_d & \bar{N}_e & \bar{N}_f \\ \bar{N}_a^T & \ddot{N}_a & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 \\ \bar{N}_b^T & {}_3\mathbf{0}^3 & \ddot{N}_b & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 \\ \bar{N}_c^T & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & \ddot{N}_c & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 \\ \bar{N}_d^T & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & \ddot{N}_d & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 \\ \bar{N}_e^T & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & \ddot{N}_e & {}_3\mathbf{0}^3 \\ \bar{N}_f^T & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & {}_3\mathbf{0}^3 & \ddot{N}_f \end{bmatrix} \begin{bmatrix} \dot{\Delta}_2 \\ \ddot{\Delta}_A \\ \ddot{\Delta}_B \\ \ddot{\Delta}_C \\ \ddot{\Delta}_D \\ \ddot{\Delta}_E \\ \ddot{\Delta}_F \end{bmatrix} = \begin{bmatrix} \dot{K}_2 \\ \ddot{K}_A \\ \ddot{K}_B \\ \ddot{K}_C \\ \ddot{K}_D \\ \ddot{K}_E \\ \ddot{K}_F \end{bmatrix}$$

where

$$\dot{N}_2 = \sum_{p=a}^{f} (\dot{B}_{2p}^T \dot{B}_{2p}) \qquad (5\times 5 \text{ submatrix})$$

$$\bar{N}_p = \dot{B}_{2p}^T \ddot{B}_{2p} \qquad (5\times 3 \text{ submatrix})$$

$$\ddot{N}_p = (\ddot{B}_{1p}^T \ddot{B}_{1p}) + (\ddot{B}_{2p}^T \ddot{B}_{2p}) \qquad (3\times 3 \text{ submatrix})$$

$$\dot{K}_2 = \sum_{p=a}^{f} (\dot{B}_{2p}^T \varepsilon_{2p}) \qquad (5\times 1 \text{ submatrix})$$

$$\ddot{K}_p = (\ddot{B}_{1p}^T \varepsilon_{1p}) + (\ddot{B}_{2p}^T \varepsilon_{2p}) \qquad (3\times 1 \text{ submatrix})$$

$${}_3\mathbf{0}^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Special methods can be used to store and solve the partitioned normal equations; however, this subject is deferred until Chap. 17.

Example 11-3

A stereopair of near-vertical photographs is taken with a 152.113-mm-focal-length camera. Photo coordinates of the images of six points in the overlap area are listed in the following table. Perform analytical relative orientation of the stereopair.

Point	Left Photo Coordinates		Right Photo Coordinates	
	x, mm	*y*, mm	*x*, mm	*y*, mm
a	–4.870	1.992	–97.920	–2.910
b	89.296	2.706	–1.485	–1.836
c	0.256	84.138	–90.906	78.980
d	90.328	83.854	–1.568	79.482
e	–4.673	–86.815	–100.064	–95.733
f	88.591	–85.269	–0.973	–94.312

Solution

1. With an ASCII text editor, create the following data file with a ".dat" extension:

```
152.113
a  -4.870    1.992   -97.920   -2.910
b  89.296    2.706    -1.485   -1.836
c   0.256   84.138   -90.906   78.980
d  90.328   83.854    -1.568   79.482
e  -4.673  -86.815  -100.064  -95.733
f  88.591  -85.269    -0.973  -94.312
```

   The first line of data in the input file is the camera focal length. The information on each of the following lines, from left to right, consists of the point identification, its *x* and *y* photo coordinates on the left photo, and its *x* and *y* photo coordinates on the right photo.

2. Run the relor program to produce the following results:

```
Exterior orientation parameters:
Parameter  Left pho  Right pho  SD right
Omega(deg)   0.0000     2.4099    0.0171
Phi(deg)     0.0000     0.5516    0.0181
Kappa(deg)   0.0000    -0.2067    0.0084
XL           0.0000    91.9740
YL           0.0000    -1.7346    0.0545
ZL         152.1130   148.3015    0.0196

Object space coordinates:
point     X        Y         Z       SD-X    SD-Y    SD-Z
    a  -4.8352   1.9730   1.0888  0.0127  0.0107  0.0975
    b  89.0970   2.7047   0.3391  0.0464  0.0109  0.0813
    c   0.2542  83.5234   1.1159  0.0117  0.0522  0.1001
    d  89.2672  82.8667   1.7862  0.0469  0.0488  0.0809
    e  -4.6333 -86.0755   1.2917  0.0126  0.0555  0.1032
    f  89.3101 -85.9635  -1.2348  0.0491  0.0528  0.0866

Photo coordinate residuals:
point   xl-res   yl-res   xr-res   yr-res
    a  -0.0001  -0.0048   0.0001   0.0047
    b   0.0001   0.0048  -0.0001  -0.0047
    c   0.0001   0.0026  -0.0001  -0.0027
    d  -0.0001  -0.0026   0.0001   0.0027
    e  -0.0000   0.0023   0.0000  -0.0022
    f   0.0000  -0.0023  -0.0000   0.0022
  RMS   0.0001   0.0034   0.0001   0.0034

Standard error of unit weight: 0.0118
Degrees of freedom: 1
```

In the upper table of output, the units of omega, phi, and kappa are degrees, while X_L, Y_L, and Z_L are in millimeters. The rightmost column lists standard deviations in the computed unknowns, also in degrees and millimeters. The middle table lists stereomodel coordinates, and their computed standard deviations, in millimeters. The lower table lists *x* and *y* photo coordinates and their residuals for both photos. Notice that, similar to the space intersection example, *x* photo coordinate residuals are nearly equal to zero for both the left and right photos, and *y* residuals are larger, nearly equal, and opposite for left and right photos. ▲

11-11 Analytical Absolute Orientation

For a small stereomodel such as that computed from one stereopair, analytical absolute orientation

can be performed using a three-dimensional conformal coordinate transformation (see Sec. C-7). This requires a minimum of two horizontal and three vertical control points, but additional control points provide redundancy, which enables a least squares solution. In the process of performing absolute orientation, stereomodel coordinates of control points are related to their three-dimensional coordinates in a ground-based system. It is important for the ground system to be a true cartesian coordinate system, such as local vertical, since the three-dimensional conformal coordinate transformation is based on straight, orthogonal axes.

Once the transformation parameters have been computed, they can be applied to the remaining stereomodel points, including the X_L, Y_L, and Z_L coordinates of the left and right photographs. This gives the coordinates of all stereomodel points in the ground system.

Example 11-4

Ground coordinates in a local vertical system for three control points are listed in the table below. For the results of the analytical relative orientation of Example 11-3, perform analytical absolute orientation using a three-dimensional conformal coordinate transformation.

Point	*X*, m	*Y*, m	*Z*, m
C	9,278.062	10,482.868	59.741
E	9,269.903	9,922.635	69.799
F	9,580.264	9,927.325	66.109

Solution

1. With an ASCII text editor, create the following data file with a ".dat" extension:

```
C    0.2542   83.5234    1.1159  9278.062  10482.868  59.741
E   -4.6333  -86.0755    1.2917  9269.903   9922.635  69.799
F   89.3101  -85.9635   -1.2348  9580.264   9927.325  66.109
#
A   -4.8352    1.9730    1.0888
B   89.0970    2.7047    0.3391
D   89.2672   82.8667    1.7862
Lpho   0   0  152.113
Rpho 91.9740 -1.7346 148.3015
#
```

   For the above input file, the first three lines relate to the control points. From left to right the data include the point identification; its *x*, *y*, and *z* stereomodel coordinates; and its *X*, *Y*, and *Z* ground coordinates. The first # sign signifies that all control has been entered, and that the data following pertain to stereomodel points whose coordinates are to be transformed into the ground system. Each data line consists of the point identification, followed by its *x*, *y*, and *z* stereomodel coordinates. The second # sign completes the data.

2. Run the "3dconf" program to produce the following results:

```
Residuals:
Point   X res   Y res  Z res
    C   0.009   0.006 -0.000
    E   0.003  -0.023  0.000
    F  -0.012   0.017 -0.000

Standard Error of Unit Weight: 0.02335

Final Results:
Param      Value     Stan.Err.
scale       3.30297   0.00015
omega      -0.9819d   0.0033d
  phi      -0.8745d   0.0061d
kappa       0.8166d   0.0026d
   Tx    9281.220     0.015
   Ty   10206.994     0.015
   Tz      60.830     0.016

Transformed Points:
Point      X          Y          Z     SDev.X SDev.Y SDev.Z
    A   9265.105   10213.339   64.073  0.015  0.015  0.017
    B   9575.295   10220.215   66.213  0.017  0.017  0.028
    D   9572.011   10485.010   66.406  0.023  0.023  0.039
 Lpho   9273.552   10215.603  563.122  0.055  0.033  0.028
 Rpho   9577.546   10214.067  555.197  0.055  0.033  0.036
```

The values listed in the top table of the output are residuals in the *X*, *Y*, and *Z* control point coordinates. The center table lists the seven parameters of the three-dimensional conformal coordinate transformation, and the lower table gives the transformed *X*, *Y*, and *Z* ground coordinates of noncontrol points, together with their computed standard deviations. ▲

References

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Eden, H. A.: "Point Transfer from One Photograph to Another," *Photogrammetric Record*, vol. 7, no. 41, 1973, p. 531.

Horn, B.K.P., "Relative Orientation," *A.I. Memo 994*, Massachusetts Institute of Technology, Cambridge, Mass., 1987.

Leupin, M. M.: "Analytical Photogrammetry: An Alternative to Terrestrial Point Determination," *Australian Surveyor*, vol. 28, no. 2, 1976, p. 73.

Maarek, A.: "Practical Numerical Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 10, 1977, p. 1295.

Smith, G. L.: "Analytical Photogrammetry Applied to Survey Point Coordination," *Australian Surveyor*, vol. 28, no. 5, 1977, p. 263.

Thompson, L. G.: "Determination of the Point Transfer Error," *Photogrammetric Engineering and Remote Sensing*, vol. 45, no. 4, 1979, p. 535.

Trinder, J. C.: "Some Remarks on Numerical Absolute Orientation," *Australian Surveyor*, vol. 23, no. 6, 1971, p. 368.

Problems

11-1. Describe two different conditions that are commonly enforced in analytical photogrammetry.

11-2. If 13 pass points are used in the analytical relative orientation of a stereopair, how many independent collinearity equations can be written?

11-3. A near-vertical aerial photograph taken with a 152.013-mm-focal-length camera contains images of five ground control points *A* through *E*. Refined photo coordinates and ground control coordinates in a local vertical system of the five points are listed in the following table. Calculate the exterior orientation parameters ω, ϕ, κ, X_L, Y_L, and Z_L for this photograph using the space resection program.

	Photo Coordinates		Ground Control Coordinates		
Point	*x*, mm	*y*, mm	*X*, m	*Y*, m	*Z*, m
A	–50.714	53.540	6,936.954	3,961.115	162.136
B	98.052	–77.430	7,792.543	3,170.883	156.891
C	–64.285	–92.908	6,840.672	3,085.747	139.923
D	7.822	–13.261	7,280.780	3,551.239	160.147
E	105.492	49.835	7,890.220	3,923.683	130.000

11-4. Repeat Prob. 11-3, except that the camera focal length is 150.023 mm, and the coordinates are as listed in the following table.

	Photo Coordinates		Ground Control Coordinates		
Point	*x*, mm	*y*, mm	*X*, m	*Y*, m	*Z*, m
A	0.387	62.053	3,260.111	4,239.222	30.333
B	–45.750	92.714	3,212.390	4,275.673	25.432
C	103.011	72.607	3,366.721	4,250.176	23.923
D	–46.703	–82.804	3,211.400	4,097.777	30.345
E	66.334	–74.520	3,323.704	4,100.444	22.222

11-5. Orientation of a stereopair of aerial photographs taken with a 150.023-mm-focal-length camera resulted in the exterior orientation values listed in the following table. If refined photo coordinates for a particular point on photo 1 are x_1 = 48.349 mm and y_1 = –30.454 mm and those for the same point on photo 2 are x_2 = –47.248 mm and y_2 = –36.097 mm, compute the object space coordinates for the point.

Parameter	Photo 1	Photo 2
ω	2.336°	3.478°
ϕ	0.923°	3.290°
κ	–1.277°	–0.001°
X_L	6,521.530 m	6,521.530 m
Y_L	4,170.999 m	4,175.844 m
Z_L	445.000 m	436.412 m

11-6. The relative orientation, found using the method described in Sec. 11-10, of a stereopair of aerial photographs taken with a 153.617-mm-focal-length camera resulted in the orientation values listed in the following table. If refined photo coordinates for a particular point on photo 1 are x_1 = 42.589 mm and y_1 = –11.113 mm and those for the same point on photo 2 are x_2 = –44.570 mm and y_2 = –9.951 mm, compute the object space coordinates for the point by space intersection.

Parameter	Right Photo
ω	1.234°
ϕ	4.321°
κ	–1.357°
X_L	100.001 mm
Y_L	–5.384 mm
Z_L	154.682 mm

11-7. A stereopair of near-vertical photographs is taken with a 149.941-mm-focal-length camera. Photo coordinates of the images of six points in the overlap area are listed in the following table. Perform analytical relative orientation on the stereopair (see Example 11-3).

	Left Photo Coordinates		Right Photo Coordinates	
Point	*x*, mm	*y*, mm	*x*, mm	*y*, mm
A	–35.217	–1.001	–112.780	–1.689
B	40.322	–6.044	–45.843	–4.819
C	100.134	4.005	11.186	6.150
D	–32.660	40.821	–111.985	36.291
E	60.400	–80.532	–25.771	–76.177
F	75.100	100.133	–14.727	96.392

11-8. Repeat Prob. 11-7 except that the camera was calibrated so its units are in pixels and focal length is 3403.77 pixels. The refined photo coordinates of the images of 12 points in the overlap area (also in pixel units) are listed in the following table.

Point	Left Photo Coordinates		Right Photo Coordinates	
	x, pixels	*y*, pixels	*x*, pixels	*y*, pixels
A	−579.0	−2,315.5	−2,548.9	−2,275.9
B	2,652.0	−2,306.2	542.8	−2,283.6
C	1,724.0	−2,298.3	−360.9	−2,272.0
D	1,741.3	2,670.1	−446.6	2,430.2
E	584.6	2,618.5	−1,505.4	2,326.6
F	344.8	2,573.9	−2,292.2	2,245.0
G	709.2	−2,316.5	−1,393.5	−2,278.4
H	2,478.7	2,212.6	294.4	2,055.0
I	−792.5	1,329.0	−2,720.9	1,123.4
J	740.8	−46.2	−1,332.0	−94.8
K	1,554.1	57.7	−715.3	24.6
L	2,269.2	−453.9	157.4	−455.4

11-9. Ground coordinates in a local vertical system for three control points *A*, *B*, and *D* are listed in the table below. Using the results of the analytical relative orientation of Prob. 11-7, perform analytical absolute orientation with the program provided (see Example 11-4), obtaining ground coordinates for the remaining points.

Point	*X*, m	*Y*, m	*Z*, m
A	4,413.674	4,598.457	197.500
B	4,604.215	4,588.008	191.861
D	4,420.136	4,702.759	199.602

11-10. Ground coordinates in a local vertical system for four control points *A*, *B*, *E*, and *F* are listed in the table below. Using the results of the analytical relative orientation of Prob. 11-8, perform analytical absolute orientation with the program provided (see Example 11-4), obtaining ground coordinates of the remaining points.

Point	*X*, m	*Y*, m	*Z*, m
A	3,207.099	5,884.630	185.629
B	4,082.690	5,838.555	212.336
E	3,589.782	7,194.489	182.169
F	3,340.175	7,207.313	157.635

CHAPTER 12
Stereoscopic Plotting Instruments

12-1 Introduction

Stereoscopic plotting instruments (commonly called *stereoplotters* or simply "plotters") are instruments designed to provide rigorously accurate solutions for object point positions from their corresponding image positions on overlapping pairs of photos. A stereoplotter is essentially a three-dimensional digitizer, capable of producing accurate X, Y, and Z object space coordinates when properly oriented and calibrated. The fact that the photos may contain varying amounts of tilt is of no consequence in the resulting accuracy; in fact, modern stereoplotters are capable of handling oblique or horizontal (terrestrial) photos. The primary uses of stereoplotters are compiling topographic maps and generating digital files of topographic information, and because these are widely practiced photogrammetric applications, the subject of stereoplotters is one of the most important in the study of photogrammetry.

The fundamental concept underlying the design of an early type of stereoplotter is illustrated in Fig. 12-1. In Fig. 12-1*a*, an overlapping pair of aerial photos is exposed. Transparencies or *diapositives*, as they are called, carefully prepared to exacting standards from the negatives, are placed in two stereoplotter projectors, as shown in Fig. 12-1*b*. This process is called *interior orientation*. With the diapositives in place, light rays are projected through them; and when rays from corresponding images on the left and right diapositives intersect below, they create a stereomodel (often simply called a *model*). In creating the intersections of corresponding light rays, the two projectors are oriented so that the diapositives bear the exact relative angular orientation to each other in the projectors that the negatives had in the camera at the instants they were exposed. The process is called *relative orientation* and creates, in miniature, a true three-dimensional stereomodel of the overlap area. After relative orientation is completed, *absolute orientation* is performed. In this process the stereomodel is brought to the desired scale and leveled with respect to a reference datum.

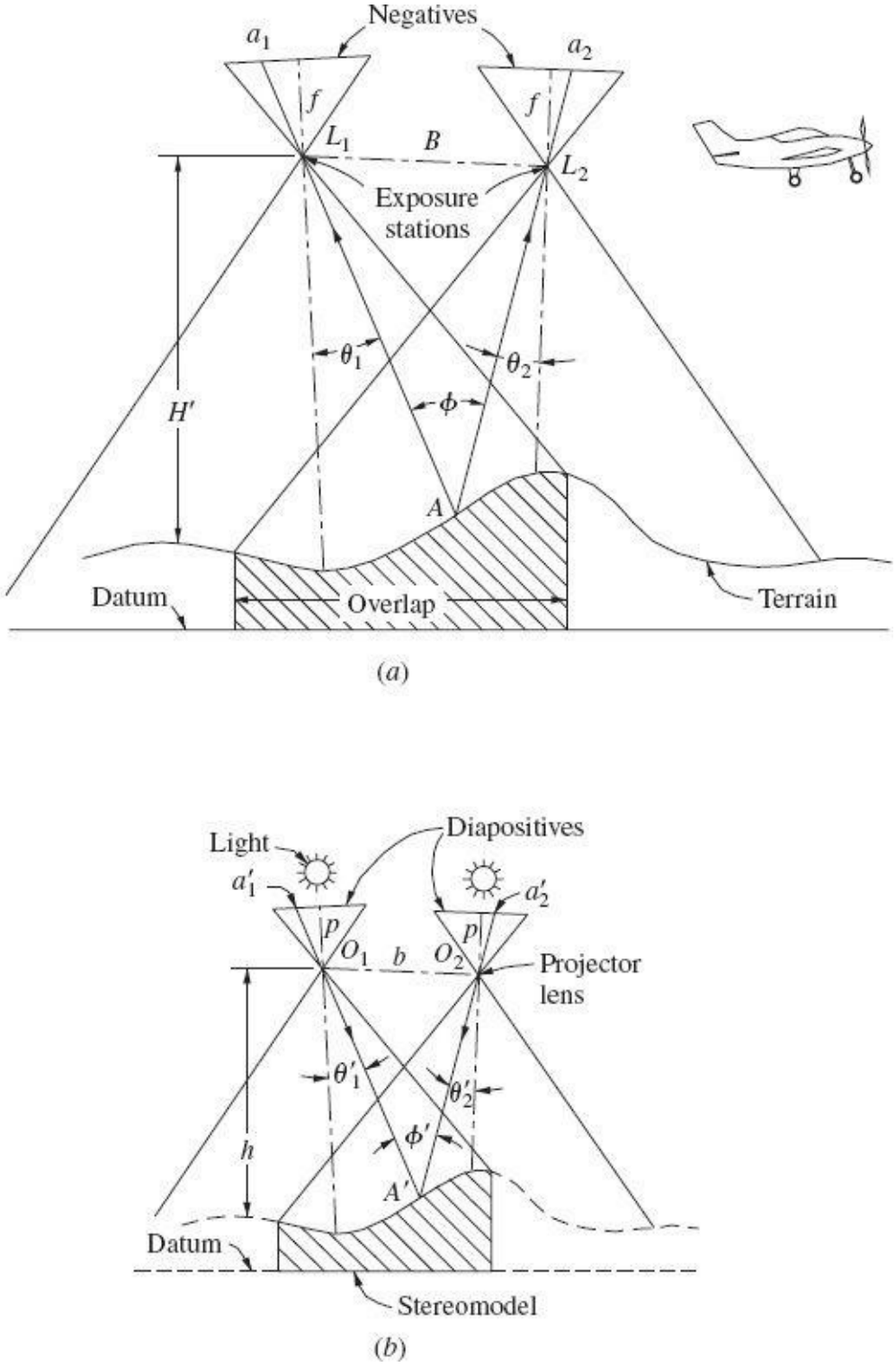


FIGURE 12-1 Fundamental concept of stereoscopic plotting instrument design. (*a*) Aerial photography; (*b*) Stereoscopic plotting instrument.

When orientation is completed, measurements of the model may be made and recorded, nowadays generally in digital, computer-compatible form. The position of any point is determined by bringing a three dimensional reference mark (the floating mark) in contact with the model point. At the position of the reference mark, the three-dimensional coordinates (X, Y, and Z) are obtained through either an analog or a digital solution. Planimetric (X, Y) positions and elevations (Z) of points are thus obtained.

12-2 Classification of Stereoscopic Plotters

A variety of stereoscopic plotting instruments have been developed over the years, each with a different design. As an aid to understanding stereoplotters, it is helpful to classify them into groups having common characteristics. This classification is divided into four general categories: (1) *direct optical projection* instruments, (2) instruments with *mechanical or optical-mechanical projection*, (3) *analytical* stereoplotters, and (4) *softcopy* stereoplotters. The first-generation stereoplotters were of direct optical projection design, and although they are only of historic interest today, a description of their operation provides a good introduction to the subject of stereoplotters. These instruments create a true three-dimensional stereomodel by projecting transparency images through projector lenses, as illustrated in Fig. 12-1. The model is formed by intersections of light rays from corresponding images of the left and right diapositives. An operator is able to view the model directly and make measurements on it by intercepting projected rays on a viewing screen (platen).

Instruments of mechanical projection or optical-mechanical projection also create a true three-dimensional model from which measurements are taken. Their method of projection, however, is a simulation of direct projection of light rays by mechanical or optical-mechanical means. An operator views the diapositives stereoscopically directly through a binocular train. While there are still some mechanical plotters in use today, details of their operation are not included in this text.

Analytical stereoplotters form a stereomodel through a purely mathematical procedure which takes place in a computer. The mathematical basis behind these plotters was introduced in Chap. 11. As with mechanical plotters, an operator views the diapositives stereoscopically directly through a binocular train. The movements of the stereoscopic images are introduced by servomotors which are under computer control. Unlike direct projection or mechanical plotters, these versatile instruments are essentially unlimited in terms of the photographic geometry they can accommodate.

Softcopy instruments are the most recent innovation in the design of stereoplotters. Fundamentally, softcopy plotters operate in the same manner as analytical stereoplotters, except that instead of viewing film (hard-copy) diapositives through binocular optics, scanned (softcopy) photographs are displayed on a computer screen and viewed directly. Special viewing systems have been designed which enable the operator to view the left image with the left eye and the right image with the right eye in order to see in stereo.

It would be difficult, if not impossible, to describe each available instrument in detail in an entire book, let alone a single chapter. For the most part, descriptions are general without reference to specific plotters and without comparisons of available instruments. To emphasize and clarify basic principles, however, examples are made of certain instruments and in some cases pictures are given. Omission of other comparable stereoplotters is not intended to imply any inferiority of these instruments. Operator's manuals which outline the details of each of the different instruments are provided by the manufacturers. Comprehension of the principles presented in this chapter should provide the background necessary for understanding these manuals.

PART I DIRECT OPTICAL PROJECTION STEREOPLOTTERS

12-3 Components

The principal components of a typical direct optical projection stereoplotter are illustrated in the schematic diagram of Fig. 12-2. The numbered parts are the (1) *main frame*, which supports the

projectors rigidly in place, thereby maintaining orientation of a stereomodel over long periods; (2) *reference table*, a large smooth surface which serves as the vertical datum to which model elevations are referenced and which also provides the surface upon which the manuscript map is compiled; (3) *tracing table*, to which the platen and tracing pencil are attached; (4) *platen*, the viewing screen which also contains the reference mark; (5) *guide rods*, which drive the illumination lamps, causing projected rays to be illuminated on the platen regardless of the area of the stereomodel being viewed; (6) *projectors;* (7) *illumination lamps;* (8) *diapositives;* (9) *leveling screws*, which may be used to tilt the projectors in absolute orientation; (10) *projector bar*, to which the projectors are attached; and (11) *tracing pencil*, which is located vertically beneath the reference mark on the platen.

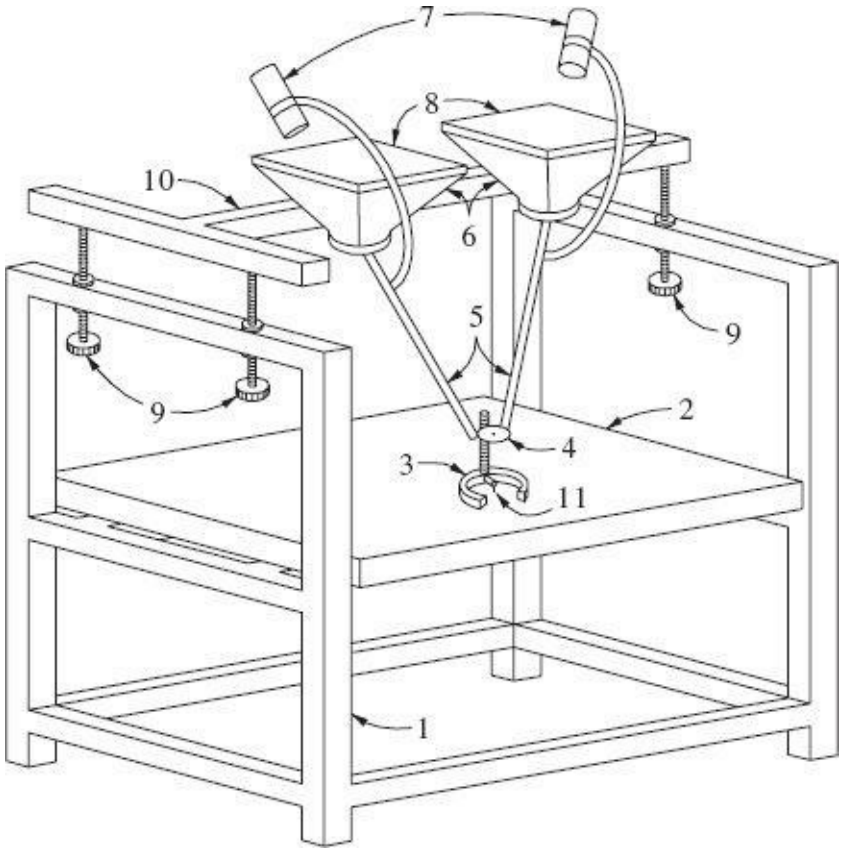


FIGURE 12-2 Principal components of a typical direct optical projection stereoplotter.

12-4 Projection Systems

In the projection systems of direct optical projection stereoplotters, diapositives of a stereopair are placed in projectors and illuminated from above. Light rays are projected through the projector objective lenses and intercepted below on the reflecting surface of the platen. The projection systems of this type of stereoplotter require that the instruments be operated in a dark room.

Stereoplotter projectors are similar to ordinary slide projectors, differing primarily in their optical precision, physical size, and capability of adjustment in angular attitude relative to one another. Since projection takes place through an objective lens, the lens formula, Eq. (2-4), must be satisfied in order to obtain a sharply focused stereomodel. In terms of the stereoplotter symbols of Fig. 12-1*b*, the lens formula is expressed as

$$\frac{1}{f'} = \frac{1}{p} + \frac{1}{h} \tag{12-1}$$

In Eq. (12-1), p is the *principal distance* of the projectors (distance from diapositive image plane to upper nodal point of the projector lens), h is the *projection distance* (distance from lower nodal point of the objective lens to the plane of optimum focus), and f' is the focal length of the projector objective lens. To obtain a clear stereomodel, intersections of projected corresponding rays must occur at a projection distance within the range of the *depth of field* of the projector lens (see Sec. 2-3).

To recreate the relative angular relationship of two photographs exactly as they were at the instants of their exposures (a process described in Sec. 12-8), it is necessary that the projectors have rotational and translational movement capabilities. These motions, six in number for each projector, are illustrated in Fig. 12-3. Three of the movements are angular rotations about each of three mutually perpendicular axes: x rotation, called *omega;* y rotation, called *phi;* and z rotation, called *kappa*. The origin of the axis system about which the rotations take place is at the projector lens, with the x axis being parallel to the projector bar. The other three movements are linear translations along each of the three axes. In general, projectors of direct optical projection stereoplotters have all three angular rotations; however, they do not necessarily have all three linear translations. As a minimum, though, they must have the x translation for changing the spacing between projectors.

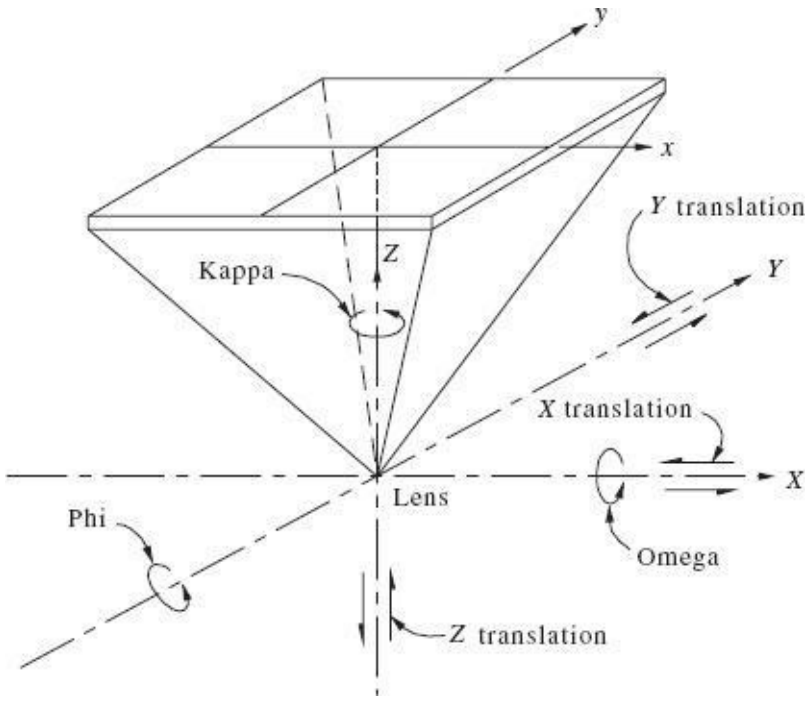


FIGURE 12-3 The six basic projector motions.

Projectors of the type illustrated in Fig. 12-2 illuminate only a small area of the diapositive at a time. This type of illumination system consists of a small, narrow-angle light source which illuminates a circular area on the diapositive approximately 4 cm in diameter. When projected through the objective lens, an area slightly larger than the platen is illuminated in the model area. Driven by means of guide rods, the lamps swing above the diapositives, following the platen and illuminating it as it moves about the stereomodel.

12-5 Viewing and Tracing Systems

The function of the viewing system of a stereoplotter is to enable the operator to view the stereomodel

three-dimensionally. Stereoviewing is made possible by forcing the left eye to view only the overlap area of the left photo while the right eye simultaneously sees only the overlap area of the right photo. The different stereoviewing systems commonly used in direct optical projection plotters are (1) the *anaglyphic* system, (2) the *stereo-image alternator* (SIA), and (3) the *polarized-platen viewing* (PPV) system.

The anaglyphic system uses filters of complementary colors, usually red and cyan (blue-green), to separate the left and right projections. Assume that a cyan filter is placed over the light source of the left projector while a red filter is placed over the right. Then, if the operator views the projected images while wearing a pair of spectacles having cyan glass over the left eye and red glass over the right eye, the stereomodel can be seen in three dimensions.

The SIA system uses synchronized shutters to achieve stereoviewing. A shutter is placed in front of each projector lens. Also, a pair of eyepiece shutters, through which the operator must look, is situated in front of the platen. The shutters are synchronized so that the left projector and left eyepiece shutters are open simultaneously while the right projector and right eyepiece shutters are closed, and vice versa. An operator therefore sees only left projector images with the left eye and right projector images with the right eye. The shutters rotate at a rapid rate so that the operator is unaware of any discontinuity in the projection.

The PPV system operates similarly to the anaglyphic system except that polarizing filters are used instead of colored filters. Filters of orthogonal polarity are placed in front of the left and right projectors, and the operator wears a pair of spectacles with corresponding filters on the left and right. In contrast to the anaglyphic system, the SIA and PPV systems both cause much less light loss, and both permit the use of color diapositives.

A system for making precise measurements of the stereomodel is essential to every stereoplotter. Measurements may be recorded as direct tracings of planimetric features and contours of elevation, or they may be taken as *X*, *Y*, and *Z* model coordinates. One of the principal elements of the measuring system of a direct optical projection stereoplotter is a tracing table. The platen (see Fig. 12-2), which can be raised or lowered, contains a reference mark in its center, usually a tiny speck of light. The reference mark appears to float above the stereomodel if the platen is above the terrain; hence it is called the *floating mark*. Vertical movement of the platen is geared to a dial, and by varying gear combinations the dial can be made to display elevations directly in meters or feet for varying model scales.

A manuscript map, preferably of stable base material, is placed on top of the reference table. The tracing table rests on the manuscript and is moved about manually in the *X* and *Y* directions. To plot the position of any point, the platen is adjusted in *X*, *Y*, and *Z* until the floating mark appears to rest exactly on the desired point in the model. A pencil point which is vertically beneath the floating mark is then lowered to record the planimetric position of the point on the map, and its elevation is read directly from the dial.

To trace a planimetric feature such as a creek, the pencil is lowered to the map and the tracing table is moved in the *XY* plane while the platen is moved up or down to keep the floating mark in contact with the stream. The pencil thereby records a continuous trace of the feature. Contours of elevation may also be traced by locking the dial at the elevation of the desired contour and then moving the tracing table about, keeping the floating mark in contact with the terrain. Tracing contours with a stereoplotter is a skill that takes years of practice to master.

12-6 Interior Orientation

As mentioned earlier, three steps are required to orient a stereoscopic plotter. The first, *interior*

orientation, includes preparations necessary to recreate the geometry of the projected rays to duplicate exactly the geometry of the original photos; e.g., angles θ_1' and θ_2' of Fig. 12-1*b* must be exactly equal to angles θ_1' and $\theta_2'$of Fig. 12-1*a*. This is necessary to obtain a true stereomodel. Procedures involved in interior orientation are (1) preparation of diapositives, (2) compensation for image distortions, (3) centering of diapositives in the projectors, and (4) setting off the proper principal distance in the projectors.

Diapositives are transparencies prepared on optically flat glass for the stereoplotter of Fig. 12-2. They are made by *direct contact* printing so their principal distances will be exactly equal to the focal length of the taking camera. Contact printing creates true geometry as long as the principal distances of the projectors are set equal to the focal length of the taking camera.

In direct optical projection plotters, compensation for symmetric radial distortion of the lens of the taking camera may be accomplished in one of the following three ways: (1) elimination of the distortion with a "correction plate" (2) varying the projector principal distance by means of a cam, and (3) use of a projector lens whose distortion characteristics negate the camera's distortion.

Each diapositive must be centered in its projector so that the principal point is on the optical axis of the projector lens. Although this problem is solved slightly differently for each instrument, it is basically done by aligning fiducial marks of the diapositive with four calibrated collimation marks whose intersection locates the optical axis of the projector.

The final step in interior orientation is to set the diapositive principal distance on the projectors. The principal distance is adjusted by either graduated screws or a graduated ring to raise or lower the diapositive image plane.

12-7 Relative Orientation

Imagine the camera frozen in space at the instants of exposure of two photographs of a stereopair. The two negatives in the camera would then bear a definite position and attitude relationship relative to each other. In relative orientation, this relative position and attitude relationship is recreated for the two diapositives by means of movements imparted to the projectors.

The condition that is fulfilled in relative orientation is that each model point and the two projection centers form a plane in miniature just like the plane that existed for the corresponding ground point and the two exposure stations. This condition is illustrated by corresponding planes $A'O_1O_2$ and AL_1L_2 of Fig. 12-1. Also, parallactic angle Φ' for any point of the stereomodel of Fig. 12-1 must equal the point's original parallactic angle Φ. The implication of the foregoing condition of relative orientation is that projected rays of all corresponding points on the left and right diapositives must intersect at a point, and this is the basis of the systematic relative orientation procedures described below.

Since relative orientation is unknown at the start, the two projectors are first positioned relative to each other by estimation. Usually, if near-vertical photography is being used, the projectors are set so that the diapositives are nearly level and so that their x axes lie along a common line. Also, the projectors are adjusted so that their Y and Z settings (distances from the projector bar in the Y and Z directions) are equal. Then, by an iterative procedure involving incremental projector adjustments while creating clear stereo views at an arrangement of spots in the overlap area, the projectors eventually become properly aligned. When finished, the projectors are in the same relative orientation as the camera stations were when the photography was acquired. The operator can then "roam" about the overlap area and see the model comfortably in three dimensions.

12-8 Absolute Orientation

After relative orientation is completed, a true three-dimensional model of the terrain exists. Although the horizontal and vertical scales of the model are equal, that scale is unknown and must be fixed at the desired value. Also the model is not yet level with respect to a vertical datum. Selecting model scale and fixing the model at that scale, and leveling the model are the purposes of absolute orientation.

Model scale is fixed within certain limits by the scale of the photography and by the characteristics of the particular stereoplotter. By comparing the geometry of Fig. 12-1*a* and *b*, model scale is seen to be the ratio of the sizes of triangles AL_1L_2 and $A'O_1O_2$. Equating these similar triangles, model scale may be expressed as

$$S_m = \frac{b}{B} = \frac{h}{H'} \tag{12-2}$$

In Eq. (12-2), S_m is model scale, b is the model air base, B is the photographic air base, h is plotter projection distance, and H' is the flying height above ground. From Eq. (12-2) it can be seen that model scale is directly proportional to model air base. Thus by varying the model air base, the model scale can be set to the desired value. However, due to limits on the range of h values, there will be corresponding limits on the range of the model air base, b.

If a preliminary model base is calculated from Eq. (12-2) and the projectors are set accordingly, then after relative orientation the stereomodel will be near the required scale. As defined by Eq. (12-2) and shown in Fig. 12-4, model scale is changed by varying the model base. If the Y and Z settings of the two projectors are equal, then model base is composed only of an X component called b_x, and model scale is varied by simply changing the model base by Δb_x, as shown in Fig. 12-4*a*.

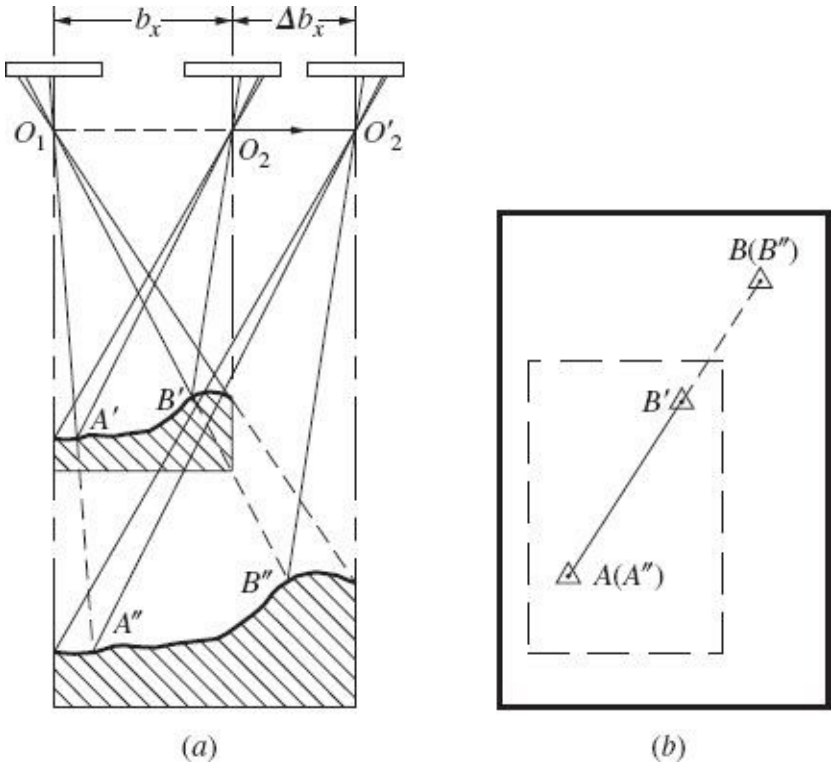


FIGURE 12-4 Changing model scale by adjusting model base. (*a*) Cross-sectional view; (*b*) plan view.

A minimum of two horizontal control points is required to scale a stereomodel. These points are plotted at adopted model scale on the manuscript map as points *A* and *B* of Fig. 12-4*b*. The manuscript is then positioned under the model, and with the floating mark set on one model point, such as *A*′, the manuscript is moved until map point *A* is directly under the plotting pencil. The floating mark is then set on model point *B*′. While the manuscript is held firmly with a fingertip at point *A*, it is rotated until map line *AB* is collinear with model line *A*′*B*′. If model line *A*′*B*′ is shorter than map line *AB*, as is the case in Fig. 12-4*b*, model scale is too small and must be increased by increasing the model base until new model line *A*″*B*″ is equal in length to map line *AB*. The model base may be set to the required value by trial and error, or Δb_x may be calculated directly from

$$\Delta b_x = b_x\left(\frac{AB}{A'B'} - 1\right) \tag{12-3}$$

In Eq. (12-3), *AB* and *A*′*B*′ are scaled from the manuscript in any convenient units. If the algebraic sign of Δb_x is negative, model scale is too large and b_x must be reduced by Δb_x. Once the model is scaled, it is recommended that a third horizontal control point be checked to guard against possible mistakes.

The final step in absolute orientation is to level the model. This procedure requires a minimum of three vertical control points distributed in the model so that they form a large triangle. As a practical matter, four points, one near each corner of the model, should be used. A fifth point near the center of the model is also desirable.

A model with a vertical control point near each corner that has not yet been leveled is shown in Fig. 12-5. Note that there are two components of tilt in the model, an *X* component (also called $\boldsymbol{\Omega}$) and a *Y* component (also called $\boldsymbol{\Phi}$). The amount by which the model is out of level in each of these components is determined by reading model elevations of the vertical control points and comparing them with their known values.

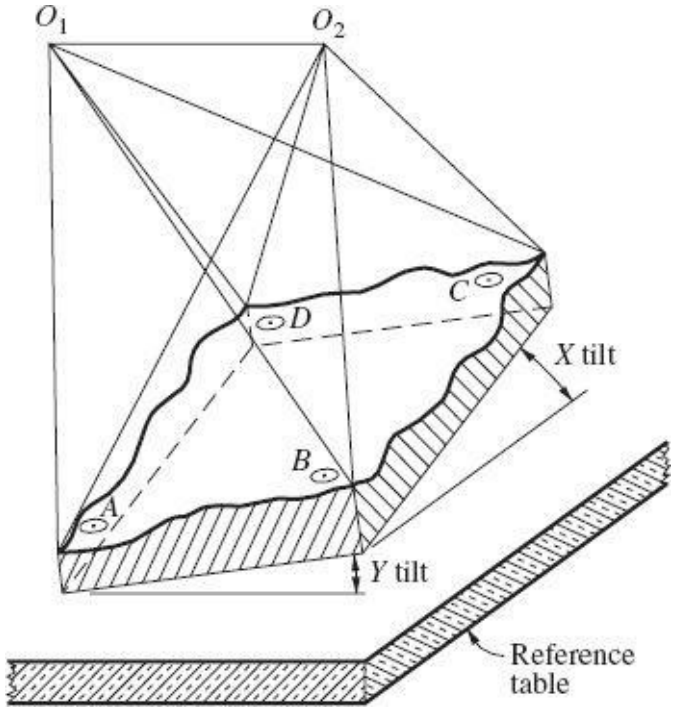


FIGURE 12-5 Stereomodel that is not level (note *X* and *Y* components of tilt).

There are different methods available for introducing the corrective *X* and *Y* tilts, depending upon the particular design of the stereoplotter. With stereoplotters such as that of Fig. 12-2, corrective tilts can be introduced by turning the leveling screws to rotate the projector bar as illustrated in Fig. 12-6. After the leveling procedure has been completed, the model will appear as it does in Fig. 12-6*b* and *d*.

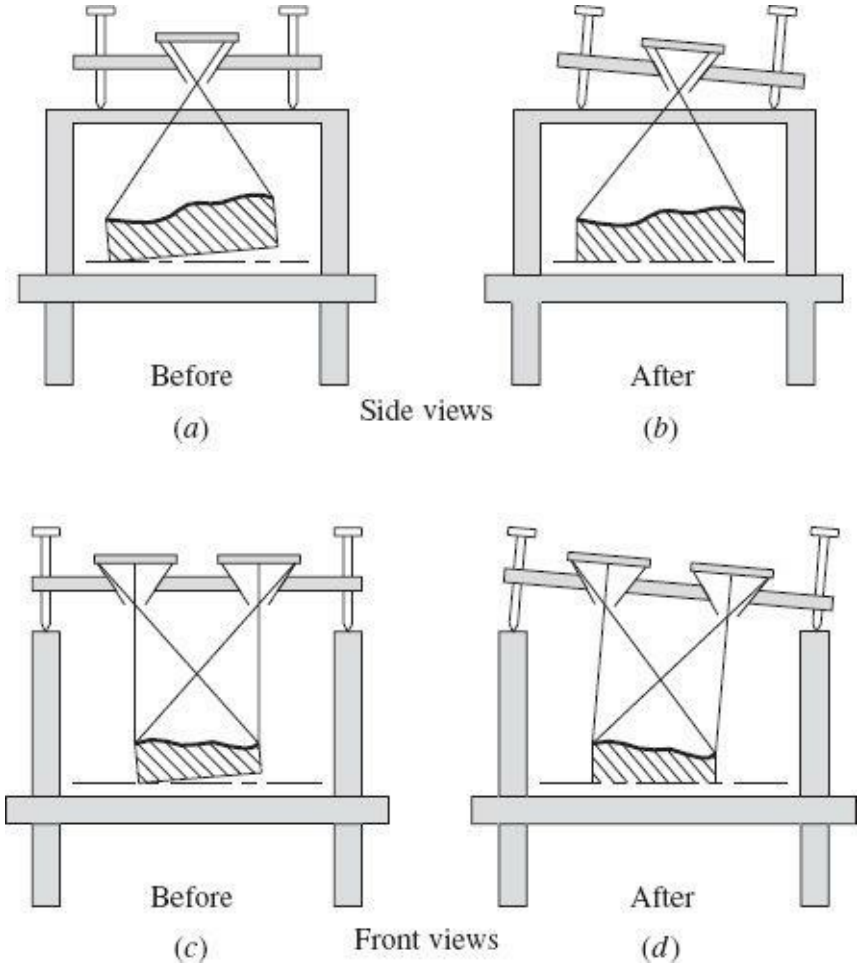


FIGURE 12-6 (*a*) and (*b*) Correcting *X* tilt of a model by *X* tilt of projector bar. (*c*) and (*d*) Correcting *Y* tilt of a model by *Y* tilt of projector bar.

The leveling operation will disturb the previously established model scale, especially if large corrective tilts are required. Also, it is likely that absolute orientation will slightly upset relative orientation. Therefore it is not practical to labor at great lengths with either relative or absolute orientation the first time through. Rather, quick orientations should be performed at first, followed by careful refinements the second time through. When orientation is completed, the manuscript map should be firmly secured to the reference table in preparation for map compilation.

PART II ANALYTICAL PLOTTERS

12-9 Introduction

Analytical plotters development was made possible by advances in computers, digital encoders, and servosystems. By combining computerized control with precision optical and mechanical components,

analytical plotters enable exact mathematical calculations to define the nature of the stereomodel. They are also easily interfaced with computer-aided drafting (CAD) systems, which facilitates map editing and updates. These instruments, with their digital output capability, are ideal for compiling data for use in geographic information systems.

Because they have no optical or mechanical limitations in the formation of their mathematical models, analytical plotters have great versatility. They can handle any type of photography, including vertical, tilted, low oblique, convergent, high oblique, panoramic, and terrestrial photos. In addition, they can accommodate photography from any focal-length camera, and in fact they can simultaneously use two photos of different focal lengths to form a model.

In comparison with analog plotters, analytical plotters can provide results of superior accuracy for three fundamental reasons. First, because they do not form model points by intersecting projected light rays or mechanical components optical and mechanical errors from these sources are not introduced. Second, analytical plotters can effectively correct for any combination of systematic errors caused by camera lens distortions, film shrinkage or expansion, atmospheric refraction, and, if necessary, earth curvature. Third, in almost every phase of their operation, they can take advantage of redundant observations and incorporate the method of least squares into the solution of the equations.

The essential capabilities of an analytical plotter are (1) to precisely measure x and y photo coordinates on both photos of a stereopair and (2) to accurately move to defined x and y photo locations. These operations are carried out under direct computer control. Digital encoders provide the ability to measure x and y photo coordinates, with the output from the encoders being read by the computer. Servomotors, which respond to signals from the controlling computer, allow the photographs to be moved to the defined locations.

12-10 System Components and Method of Operation

An analytical plotter incorporates a pair of precision comparators which hold the left and right photos of a stereopair. Figure 12-7 shows a schematic diagram of a single comparator which employs a lead screw for measurement and movement. The plate carrier, which holds the photograph, moves independently in the X and Y directions as the servomotors turn the corresponding lead screws. Digital encoders attached to the shafts of the motors sense the position of a point on the photo relative to an index mark (half mark) which is fixed.

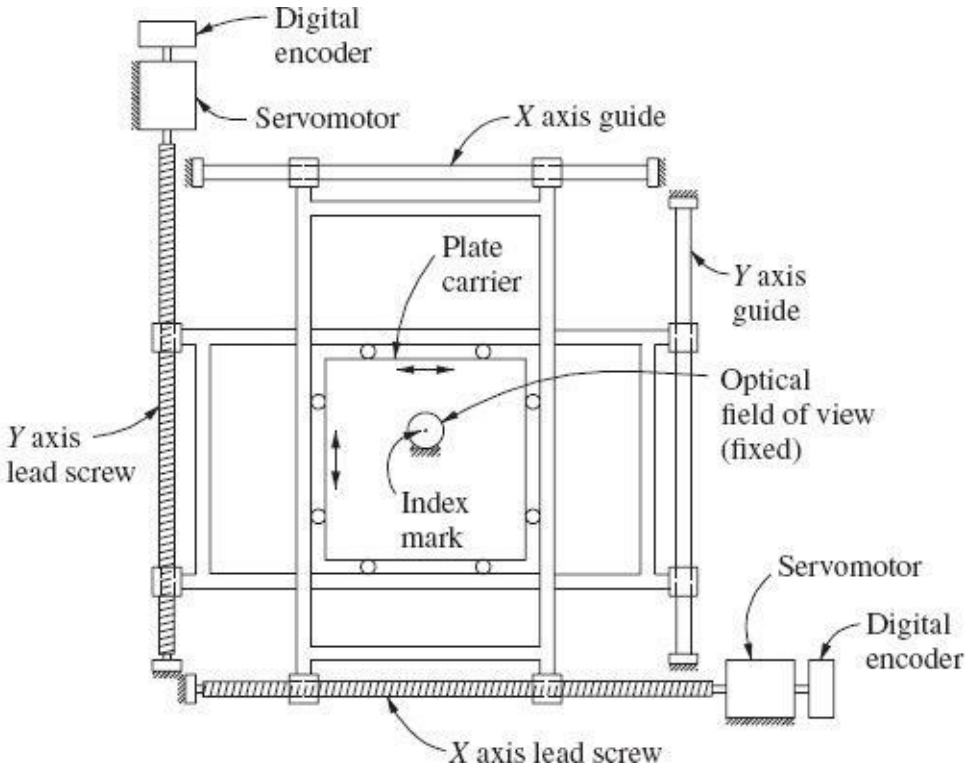


FIGURE 12-7 Schematic representation of a comparator from an analytical plotter.

Figure 12-8 illustrates how the primary components of an analytical stereoplotter interact. The heart of the system is the *controller computer* which interfaces with the *operator controls*, *servomotors*, *encoders*, and the *CAD computer*. The controller computer accepts input from the operator controls and calculates left and right plate positions from these inputs. It then operates the servomotors to move the plates, stopping when the encoders indicate that the correct positions have been reached.

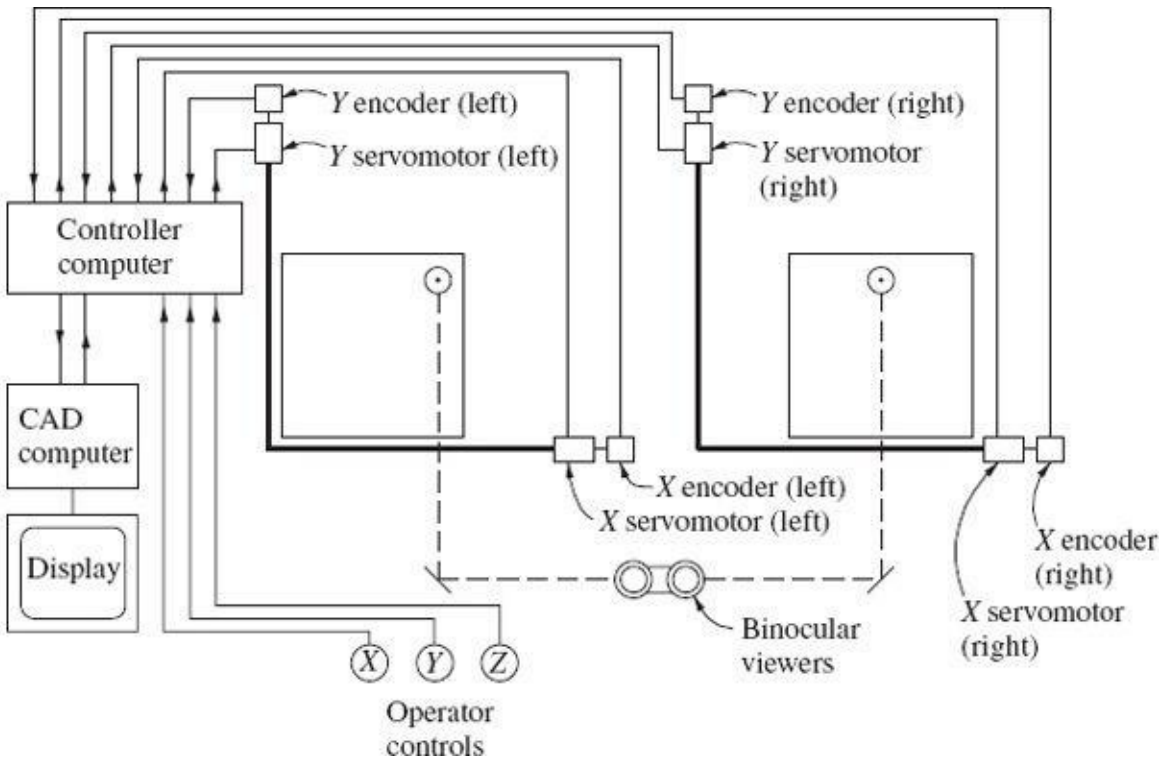


FIGURE 12-8 Schematic diagram of components and operation of an analytical plotter.

Analytical plotters form neither an *optical* nor a *mechanical* model, rather, they compute a *mathematical* model based on the principles of analytical photogrammetry presented in Chap. 11. This mathematical model is established through numerical versions of interior, relative, and absolute orientation. Although the exact methods used to accomplish these orientations may vary among instruments, the fundamental approach is the same. Typically, the orientation software is a distinct module, separate from the data collection (mapping) software. This allows flexibility in the choice of available data collection software for a given analytical plotter.

Although the development of the softcopy stereoplotter has reduced the demand for analytical plotters and major manufacturers are no longer producing new plotters they are still widely used. Two popular analytical stereoplotters are the Zeiss P-3, and the Leica SD3000, shown in Figs. 12-9 and 12-10, respectively.

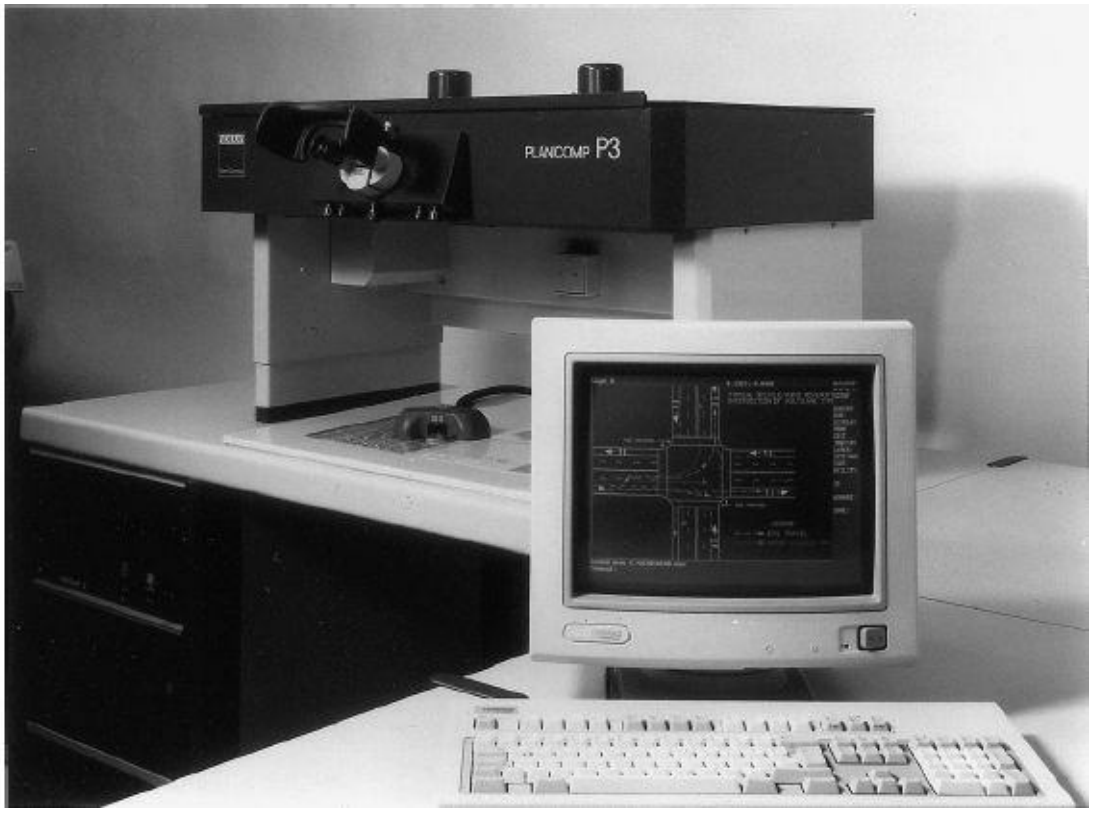


FIGURE 12-9 Zeiss P-3 analytical plotter. (*Courtesy Carl Zeiss, Inc.*)

FIGURE 12-10 Leica SD3000 analytical plotter. (*Courtesy LH Systems, LLC.*)

12-11 Analytical Plotter Orientation

As is necessary with all analog stereoplotters, interior, relative, and absolute orientation are also required for analytical plotters prior to going into most modes of operation. In all phases of orientation, and operation for that matter, a dialogue is maintained between operator and instrument. The system's computer guides the operator through each of the various steps of operation, calling for data when needed and giving the operator opportunities to make certain decisions.

Prior to using an analytical plotter, its measuring system should be calibrated using a precise grid plate. This plate consists of a flat piece of glass upon which an array of fine crosses has been etched. The coordinates of the crosses on a high-quality grid plate are known to within 1 or 2 μm. The grid plate is placed on the plate carrier, and the coordinates of each of the crosses are measured. A two-dimensional coordinate transformation is then computed which relates the raw plate coordinates of the analytical plotter to the accurate xy system of the grid plate. An analytical plotter should be recalibrated approximately every year as part of routine maintenance or anytime the instrument is moved, to ensure accurate results.

The orientation and operation of all analytical plotters are quite similar. The following subsections describe, in general, the three orientation steps for analytical plotters.

12-11-1 Interior Orientation

As mentioned in Sec. 12-6, interior orientation consists of (1) preparation of diapositives, (2) compensation for image distortions, (3) centering of diapositives, and (4) setting off the proper

principal distance. Steps 2 and 4 simply amount to accessing the appropriate lens distortion coefficients and camera focal length from a data file. By using these parameters in the appropriate photogrammetric equations, these steps of interior orientation are readily accomplished. Diapositives used with analytical plotters are typically contact prints made on film transparencies, although some plotters can accommodate paper (opaque) prints. The individual diapositives of a stereopair are placed on the left and right plate carriers and held in place by glass covers. Centering of the diapositives is accomplished by measuring the *X* and *Y* plate coordinates (based on the encoders) of the fiducials of each photo. This phase of operation is aided by computer-activated servomotors, which automatically drive the measuring mark to the vicinity of the fiducials. A fine pointing is then made by the operator. As few as two fiducials can be measured, but more are recommended and up to eight should be measured if they are available, to increase redundancy. Individual two-dimensional coordinate transformations (see App. C) are then calculated for the left and right photos. This establishes the relationship between the *XY* plate coordinates and the *xy* calibrated fiducial coordinates, and at the same time it compensates for film shrinkage or expansion. A choice of coordinate transformations is available, but usually the affine or projective types are used (see Secs. C-6 and C-9). Residuals for this solution will be displayed so that the operator can either accept them or remeasure one or more fiducials. When the solution is accepted, the interior orientation parameters are stored in the computer.

12-11-2 Relative Orientation

Relative orientation is achieved by numerical techniques as described in Sec. 11-10. The operator measures the left and right photo coordinates of at least six pass points to provide observations for the least squares relative orientation solution. In areas of imagery that contain discrete features such as manhole covers and sidewalk intersections, the operator sets the individual left and right index marks on the images of the feature in each photo to measure their coordinates. In open areas such as grass-covered fields, the operator adjusts the floating mark while viewing in stereo until it appears to rest exactly on the ground. Once the raw plate coordinates are observed for all pass points, they are transformed to the calibrated fiducial system, and then lens distortion and atmospheric refraction corrections are applied, resulting in refined coordinates. Numerical relative orientation is then calculated, resulting in values for the exterior orientation parameters ($\boldsymbol{\omega}$, $\boldsymbol{\Phi}$, $\boldsymbol{\kappa}$, X_L, Y_L, Z_L) for both photos. Again the residuals will be displayed, and the operator has the option of discarding certain points, adding others, or accepting the solution. When relative orientation is accepted, the orientation parameters are stored in the computer. At this point the operation of the analytical plotter changes from a two-dimensional to a three-dimensional mode. The *X*, *Y*, and *Z* operator controls provide model coordinate input to the controller computer which then drives the plates to the appropriate image locations.

12-11-3 Absolute Orientation

Absolute orientation is commonly achieved by a three-dimensional conformal coordinate transformation. While viewing in stereo, the operator places the floating mark on the images of the ground control points which appear in the model and records the *XYZ* model coordinates for each. The computer then accesses a previously established data file which contains the ground coordinates for the control points, thus enabling a three-dimensional conformal coordinate transformation to be calculated. For absolute orientation a minimum of two horizontal and three well-distributed vertical control points are required, just as for absolutely orienting analog plotters. More than the minimum is recommended, however, so that a least squares solution can be made. Once again, residuals are displayed, and the operator has the option of discarding or remeasuring individual points. After the

solution is accepted, the transformation parameters are stored, the analytical stereoplotter is fully oriented, and it is ready to be used as a three-dimensional mapping tool.

12-12 Three-Dimensional Operation of Analytical Plotters

After the steps of orientation are complete, the parameters from each step are available for real-time control of the plotter. The operator supplies input of *X*, *Y*, and *Z* ground coordinates through hand (and/or foot) controls, and the instrument responds by calculating theoretical left and right plate coordinates for this ground position. The servomotors then move the left and right plates to these coordinates so that the floating mark appears at the corresponding location in the model. At first glance, this sequence of operation seems to be out of order. Users of analytical plotters often incorrectly assume that their hand control inputs directly move the index marks to positions in the photos and that ground *XYZ* coordinates are then computed. This confusion may arise due to the real-time nature of plotter operation.

The top-down flow of this real-time operation is illustrated in Fig. 12-11. In this figure, computational steps are indicated with solid boxes, inputs represented with dashed boxes, and instrument movements indicated with ellipses. The process begins when the operator adjusts the *XYZ* controls which provide the primary inputs. The controlling computer then calculates the corresponding *XYZ* model coordinates by applying parameters of the three-dimensional conformal coordinate transformation, which was calculated in the absolute orientation step. Note that the three-dimensional transformation was computed going from ground system to model.

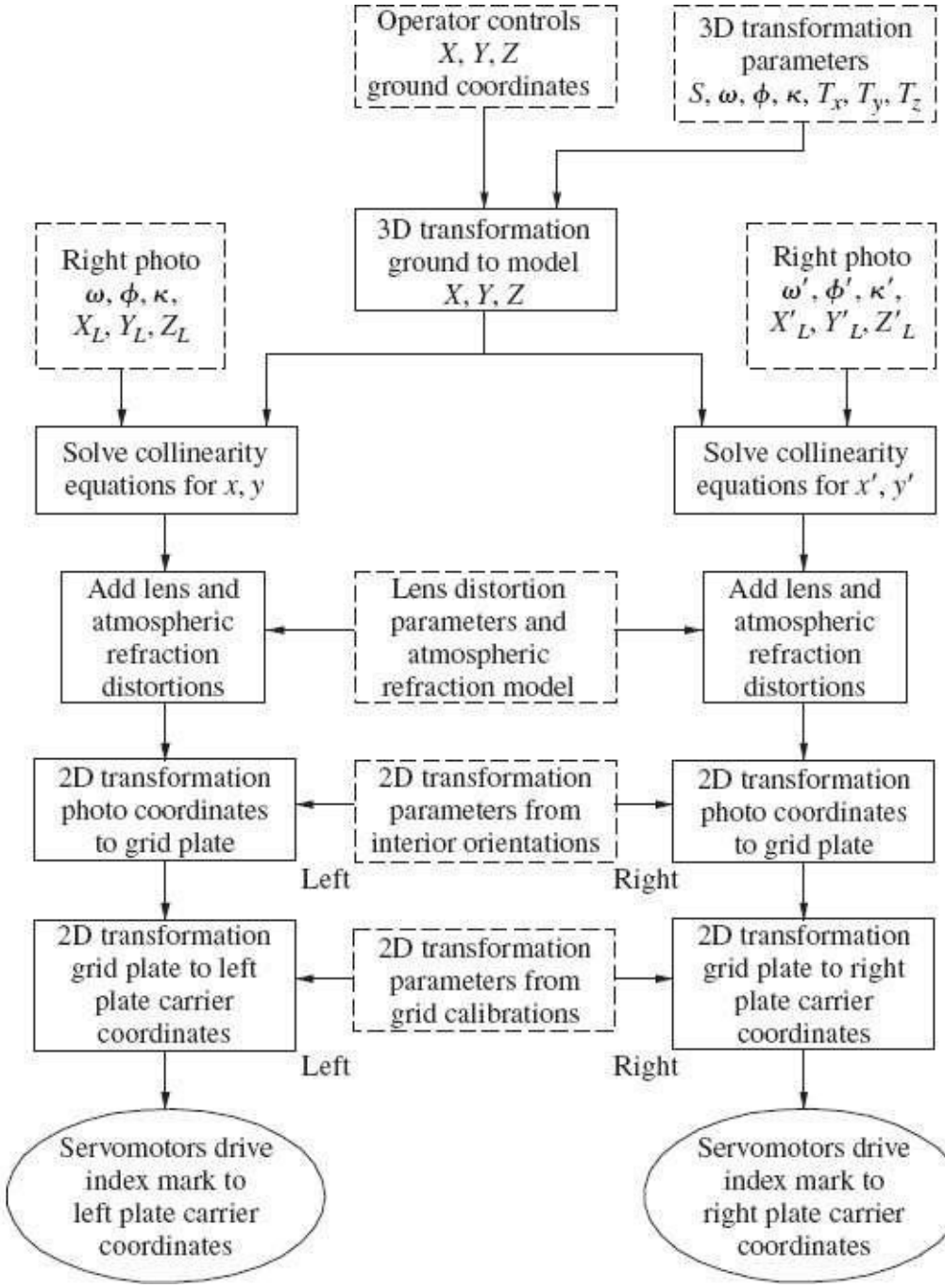


FIGURE 12-11 Flowchart showing the real-time, three-dimensional operation of an analytical stereoplotter.

After the *XYZ* model coordinates have been computed, the computer solves the collinearity equations [see Eqs. (11-1) and (11-2)] for x and y coordinates on the left photo and x' and y' coordinates on the right. The calculations are made using the *XYZ* model coordinates as well as the exterior orientation parameters $\omega, \phi, \kappa, X_L, Y_L$, and Z_L of the left photo and $\omega', \phi', \kappa', X_L', Y_L'$, and Z_L' of the right. The photo coordinates directly computed from the collinearity equations contain no distortions and may therefore be considered ideal coordinates. The actual images in the photographs, however, do not occupy these ideal positions due to lens distortion and atmospheric refraction. Therefore, when the servomotors move the plates, in order for the actual images in the photos to correctly coincide with the index marks in the binoculars, distortions must be added to the ideal coordinates computed by the collinearity equations.

The next step is to transform from the photo coordinate system to the calibrated grid plate coordinate system. This calculation is based on the two-dimensional coordinate transformation

parameters which were computed during interior orientation. The final step is to transform from grid plate coordinates to the plate carrier coordinate system. This calculation relies on the two-dimensional coordinate transformation parameters which were computed during grid plate calibration. With the plate carrier coordinates known for both photographs, the servomotors are instructed to drive the left and right index marks to these locations. All these calculations can be performed in a fraction of a millisecond on a modern computer, giving a real-time response. The only lag in the system is due to the finite speed of the servomotors. However, unless fast, abrupt movements are made with the operator controls, the entire process happens smoothly and with instantaneous response.

The *XYZ* stereoplotter controls are directly linked to the CAD system inputs. The operator could, in fact, draw on the CAD system using these controls without looking into the plotter, although the drawn entities would have no relationship to the ground if this were done. By looking into the plotter, the operator can see the movement of the floating mark as it responds to *XYZ* input. The operator guides the floating mark in *X*, *Y*, and *Z* until it appears to rest directly on the feature of interest. At that point, the final position of the operator controls defines the *XYZ* ground position of the feature, and it is then recorded into the CAD system.

12-13 Modes of Use of Analytical Plotters

The modes of use of analytical plotters vary only slightly with the different instruments. Generally all can perform planimetric and topographic mapping. In this case the operator simply traces out features and contours by keeping the floating mark in contact with each feature. The ground coordinates from the operator controls are directly transmitted to the CAD computer for subsequent off-line editing and plotting, or for the development of digital elevation models. “Layered information” needed for the databases of geographic information systems can also be conveniently compiled and stored.

Analytical plotters can also be operated in a profiling mode. In this operation *X*, *Y*, and *Z* coordinates can be measured along predetermined lines in the model. The locations and directions of desired profile lines can be input to the computer in terms of *XY* ground coordinates, and the machine will automatically drive the measuring mark to the corresponding locations while an operator monitors the floating mark to keep it in contact with the ground. The *XYZ* plotter output is continuously recorded by the CAD computer. Profile data can be used directly for various engineering applications such as investigating earthwork quantities along proposed highway centerlines, or they can be used to generate terrain elevation data for subsequent digital orthophoto production (see Sec. 13-7 and 13-8).

Another common application of analytical plotters is in aerotriangulation. Here various modes of operation are possible, from independent model triangulation to simultaneous bundle adjustment (see Chap. 17). When the aerotriangulation is completed, the diapositives can be used in other stereoplotters or photogrammetric scanners, or they can be reset in the analytical plotter based upon the now available ground control information. As a result of aerotriangulation, the parameters of relative and absolute orientation become known. Analytical plotters have a “reset” capability which enables reading in these known parameters, whereupon an instantaneous relative and absolute orientation can be achieved automatically, thus bypassing the measurements otherwise needed to determine those two steps. In fact, if for any reason a model that was once set should need to be reset to obtain additional data, the orientation parameters from the first setting can be recalled and the system will automatically perform relative and absolute orientation. Anytime photos are moved, however, interior orientations must be redone since the positions of the photographs on the plate carriers will have changed.

Analytical plotters can also be used simply as monocomparators or stereocomparators, where

image coordinates are measured and recorded for use in aerotriangulation. Many analytical stereoplotters contain a software module which performs aerotriangulation. Alternatively, photo coordinate data can be recorded in a computer file in various formats, suitable for any of several commercially available aerotriangulation software programs.

PART III SOFTCOPY PLOTTERS

12-14 Introduction

The latest stage in the evolution of stereoplotters is that of softcopy plotters. Advances in the computational speed, memory capacity, disk storage, and display monitors of computer systems along with accurate, high-resolution photogrammetric scanners and digital cameras have led to these fully computerized systems. The fundamental operation of a softcopy plotter is the same as that of an analytical plotter except that instead of employing servomotors and encoders for point measurement, softcopy systems rely on digital imagery. Softcopy plotters can perform all the operations of an analytical plotter and, due to their design, can perform a wealth of digital image processing routines as well.

Softcopy stereoplotters have had a tremendous impact on the practice of photogrammetry. By replacing costly optical and mechanical components with digital display manipulation, softcopy plotters have become less expensive and more versatile than analytical plotters. The core of a softcopy plotter is a set of computer software modules that perform various photogrammetric tasks. Software is the distinguishing characteristic, while associated computer hardware may be purchased “off the shelf” from a variety of vendors.

Besides the software, another essential component of a softcopy plotter is a computer with a high-resolution graphics display. The computer must be capable of manipulating large digital images efficiently and must be able to display left and right photos of a stereopair simultaneously or nearly so. At the same time, a special configuration is required so that the operator’s left eye views only the left photo and the right eye views only the right photo.

Manual use of a softcopy plotter is most similar to that of an analytical stereoplotter. Orientations can be performed by placing the measuring mark (a single pixel or small pattern of pixels on the display) on necessary points, followed by analytical calculations being performed to compute the orientation. Once oriented, the softcopy plotter can be used in the three-dimensional mode to measure (digitize) topographic features.

In addition to the manual uses available with softcopy plotters, these systems offer many automatic features not found on analytical plotters. One of the most useful automatic capabilities is the ability to perform routine point measurement by computer processing, requiring little or no operator input. This capability can significantly speed up the process of orienting the plotter, as well as assist in collecting digital elevation model information, profiles, and cross-sections. Another convenience offered by softcopy plotters is *vector superimposition*, in which topographic map features (lines, points, etc.) are superimposed on the digital photos as they are being digitized. This capability, also found on a few analytical stereoplotters, allows operators to keep their gaze in one place rather than constantly moving back and forth from photograph display to CAD display.

12-15 System Hardware

The fundamental hardware requirement for a softcopy stereoplotter is a high-performance computer workstation with advanced graphics capability. The computer must have not only substantial processing speed, but also large amounts of random access memory and mass storage (hard disk space). Many systems employ multiple processors to increase computational speed. This high speed and the large memory capacity are necessary because digital image files used in softcopy plotters may be several hundred megabytes or more. In addition, some form of archival storage such as optical disks is required for offline storage of digital images. The system also requires operator controls for *X*, *Y*, and *Z* position within the stereomodel. On a softcopy system, the *X* and *Y* controls are usually implemented as some type of computer "mouse," while the *Z* control is typically a small wheel which can be rotated by the operator's thumb.

For the computer to serve as a stereo workstation, additional hardware is required to enable the operator to see the left image with the left eye and the right image with the right eye. Various approaches are available to provide stereoviewing capability, such as polarizing filters and alternating shutters. In the polarizing filter approach, a computer, display system is configured so that light from the left image and light from the right image are orthogonally polarized. Meanwhile, the operator wears a simple pair of spectacles consisting of orthogonally polarized filters.

Figure 12-12 illustrates a stereo display configuration for the polarizing filter approach. The lower and upper displays (LCD monitors) are arranged so that the tops are nearly joined at the *hinge point* and the bottoms are spread out at an angle of approximately 110°. A partially reflective mirror bisects the angle of the displays with the reflective side facing the upper display and the anti-reflective side facing the lower display. Understanding how the stereo image is formed requires knowledge of an interesting detail regarding LCD computer monitors.

The physical construction of an LCD requires the use of polarizing filters and as a result, light emanating from an LCD monitor is polarized, usually in a 45° orientation. In the polarizing filter approach, light from the lower display passes through the anti-reflective side of the mirror without changing its polarization angle. It will then pass through the filter (left filter in Fig. 12-12), that has the same polarization angle and will be visible to the left eye. However, this polarized light will not pass through the filter that has the orthogonal orientation (right filter in Fig. 12-12). Light from the upper display changes polarization 90° by reflecting off the mirror and will therefore be blocked by the left filter but will pass through the right. The result is that the left eye will see the left image (from the lower display) and the right eye will see the right image (from the upper display). Figure 12-13 shows the polarizing filter display configured for a softcopy system. One additional detail that is addressed in the design is the fact that the reflected image of the upper display will be reversed in a left/right sense. An additional hardware device or software emulator is required to reverse the initial monitor image to that it appears normal in its mirrored state.

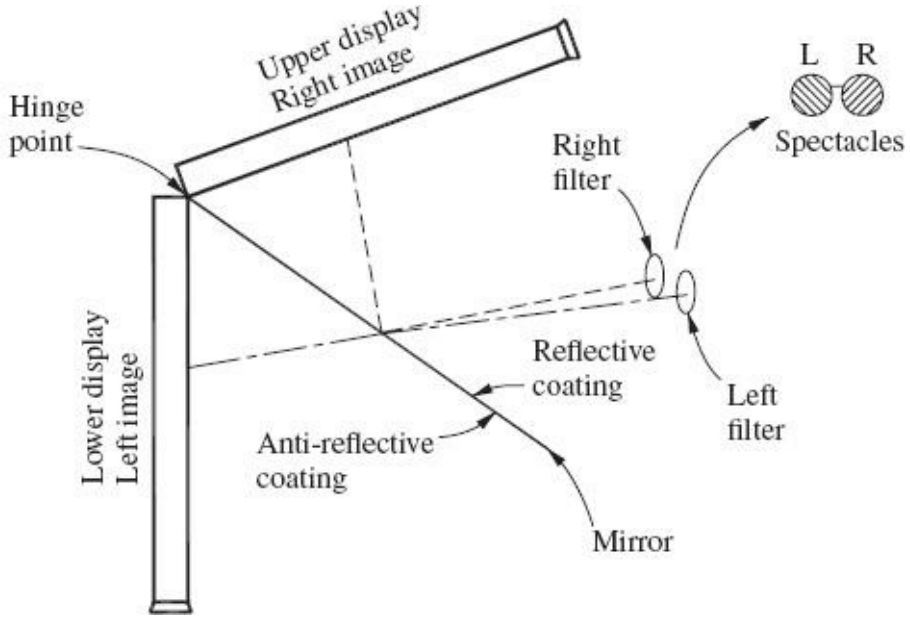


FIGURE 12-12 Stereoviewing principle of the polarizing filter screen display.

FIGURE 12-13 Softcopy photogrammetric workstation with polarizing filter stereo imaging. (*Courtesy BAE Systems*.)

A second approach to stereoviewing uses a display monitor which shows alternating left and right images at 120 Hz, along with special viewing glasses that have liquid crystal shutters (LCS) which alternate at the same rate. The LCS glasses receive signals transmitted by the computer—often from an infrared device—which control the alternating left and right masking functions. At a particular instant, the left image is displayed on the monitor, and at the same time, the LCS mask over the right eye turns opaque while the LCS mask over the left eye is clear, as illustrated in Fig.12-14*a*. Since the

right eye is blocked and the left eye is unobstructed, the operator sees the left image with the left eye. A split second later, the right image is displayed on the monitor, and at the same time the LCS mask over the left eye turns opaque while the LCS mask over the right eye is clear, as illustrated in Fig. 12-14*b*. This causes the operator to see the right image with the right eye. When this is repeated at a rate of 120 Hz, the operator is unaware of the alternating images, yet the proper stereoview is created. The system shown in Fig. 12-15, uses this method of stereo image display.

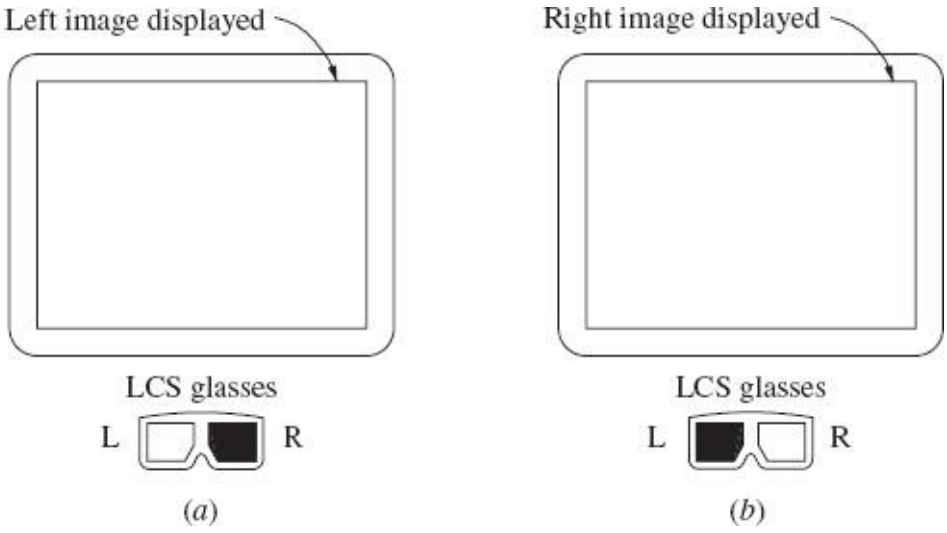


FIGURE 12-14 Stereoviewing principle of the alternating LCS shutter display. (*a*) view of the screen with the left image displayed and right LCS mask closed (*b*) view of the same screen a split second later, with the right image displayed and left LCS mask closed.

FIGURE 12-15 VR mapping workstation. (*Courtesy Cardinal Systems*.)

Each of the aforementioned methods of stereoviewing has its advantages. An advantage of the polarizing screen method is that more than one person can view the stereo display at one time, with

extra sets of inexpensive polarized spectacles. The alternating-shutter approach shares the advantage of more than one person viewing the stereo display at a time, although this requires multiple sets of LCS glasses, which are more expensive than simple polarized spectacles. In a production environment this advantage is minor, except perhaps for training novice operators. The polarizing screen and alternating-shutter methods also allow the operator more freedom of movement compared to binocular viewing systems associated with analytical plotters.

12-16 Image Measurements

As is the case with all stereoplotters, manual image measurements are accomplished through operator control of a floating mark. On a softcopy stereoplotter, a floating mark consists of left and right half marks which are superimposed on the left and right images, respectively. An individual half mark consists of a single pixel, or small pattern of pixels in the shape of a dot, cross, or more complex shape. The pixel(s) of the half mark is (are) set to brightness values and/or colors which give a high contrast with the background image. When the operator moves the *X*, *Y*, or *Z* control, the positions of the individual half marks move with respect to the background images. Once the floating mark visually coincides with a feature of interest, the operator can press a button or foot pedal to record the feature's position.

Two approaches are used for half mark movement: a fixed mark with a moving image or a fixed image with a moving mark. The first approach, which mimics the action of an analytical stereoplotter, keeps the individual half marks at the same physical position on the screen while the images pan across the screen under operator control. This implementation places a high demand on the computer display in terms of data transfer rate, thus requiring a high-speed computer graphics adapter. The second approach works by keeping the left and right images in a fixed position on the display while the individual half marks move in response to the operator's input. While this approach places far less demand on the display adapter, it creates additional problems. One problem is that when the half marks approach the edge of the image area, the images must be reloaded and redisplayed so that the half marks can be placed at the center of the viewing area. This discontinuous image shift can be quite disruptive to the operator, since it requires a corresponding shift in the direction of the operator's gaze. Another problem, which is more subtle, is that the operator's eyes must follow the floating mark as it moves, which can cause eye strain. With the fixed half mark approach, the operator's gaze remains essentially fixed while the images move around it.

A key advantage afforded by softcopy plotters is their ability to make point measurements automatically. This is accomplished through some form of pattern-matching technique, wherein a small subarray of digital numbers from the left image is matched with a corresponding subarray from the right image. Finding the matching position is equivalent to manually setting the floating mark so that it appears to rest directly on a feature in the stereomodel. Various methods are available for performing the pattern-matching operation, and some of these are discussed in Sec. 15-8.

12-17 Orientation Procedures

Softcopy plotters are oriented in a similar fashion to that of analytical plotters; i.e., the same three steps of interior, relative, and absolute orientation must be performed. The key difference between orientations of softcopy and analytical plotters is that softcopy systems allow for greater automation in the process. Interior orientation, which primarily consists of pointing on the fiducial marks, can be done directly under operator control, or by use of pattern-matching methods. Systems that use pattern

matching attempt to find the positions of the fiducial marks by matching a standard image of the fiducial, sometimes called a *template*, with a corresponding subarray from the image. Once all fiducials have been located, a two-dimensional transformation can be computed to relate image coordinates (row and column) to the fiducial axis system.

Relative orientation can also be greatly assisted by automatic pattern matching. Small subarrays from the left image in the standard pass point locations are matched with corresponding subarrays from the right image. Once a sufficient number of pass points have been matched (generally at least six points), a relative orientation can be computed. Accuracy of the relative orientation can be improved by matching additional pass points, thus providing greater redundancy in the least squares solution.

Absolute orientation is much less amenable to automation than interior or relative orientation. In absolute orientation, three-dimensional measurements must be made on the positions of ground control points which appear in the model. Since ground control points can have varying shapes and can appear anywhere within the model, they are more difficult to locate by automatic pattern-matching techniques. In these cases, manual pointing on the control points is usually done.

One situation where absolute orientation can be automated occurs when a block aerotriangulation has previously been performed on the digital images. As a result of the aerotriangulation, exterior orientation parameters will have been determined for each photo. Having known exterior orientation parameters essentially defines the absolute orientation, thus no additional measurements need to be taken.

12-18 Epipolar Geometry

As mentioned in Sec. 15-3, automatic methods can be used to set the floating mark on a feature. This involves pattern matching, which can be a computationally intensive task. To reduce the computational burden, it is helpful to constrain the search area so as to avoid unnecessary calculations. By exploiting the principle of *epipolar* geometry, search regions can be constrained along a single line.

As described in Sec. 11-5, coplanarity is the condition in which the left and right camera stations, an object point, and the left and right images of the object point lie in a common plane. If relative orientation is known for a given stereopair, the coplanarity condition can be used to define epipolar lines. This situation is illustrated in Fig. 12-16. The figure shows the intersection of the *epipolar plane* (any plane containing the two exposure stations and an object point, in this instance plane L_1AL_2) with the left and right photo planes. The resulting lines of intersection are the *epipolar lines*. They are important because, given the left photo location of image a_1, its corresponding point a_2 on the right photo is known to lie along the right epipolar line. Based solely upon the location of image point a_1, object point A could be located anywhere along line L_1A, at an arbitrary Z elevation. For example, if object point A were located at position A', its image would still be at point a_1 in the left photo. Based on an assumed object point position of A', the corresponding location of its image on the right photo, a'_2 can be calculated by the collinearity equations (see App. D). A small subarray of digital numbers at the location of point a'_2 can be compared with a corresponding subarray at the location of point a_1. Since these two locations do not both correspond to images of object point A, the patterns would not match, and another point on the right epipolar line would be tried. The search continues along the epipolar line until it zeros in on the corresponding image at a_2, where the patterns match. The coordinates of image points a_1 and a_2 can then be used to determine the three-dimensional object space coordinates of point A.

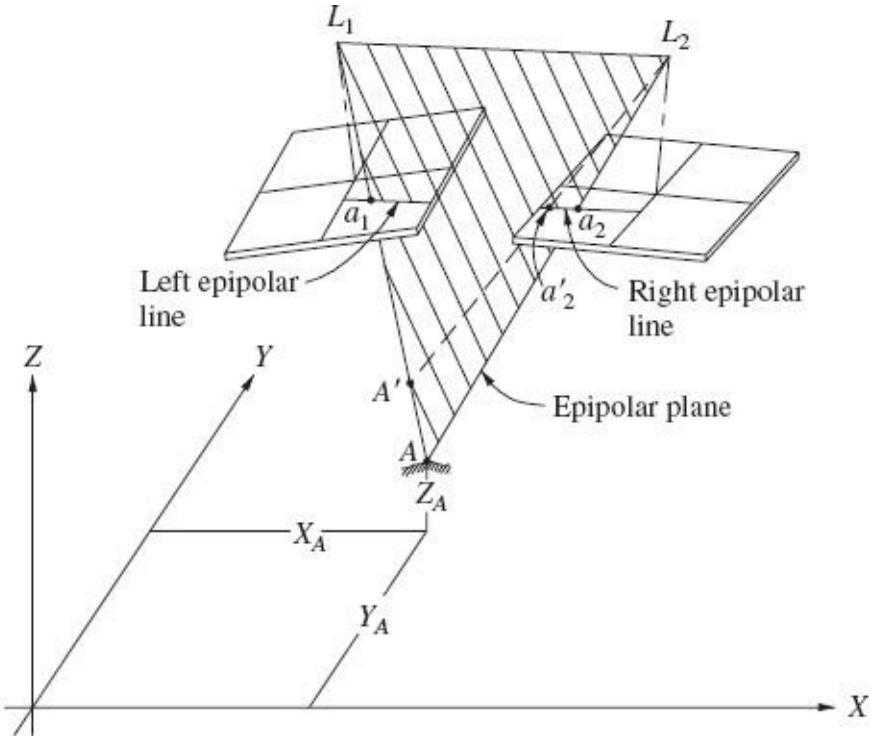


FIGURE 12-16 Epipolar geometry of a stereopair of photos.

By using the principle of epipolar geometry, searching for matching points is much more efficient. This is especially important when a large number of points must be matched, such as when a digital elevation model is created (see Sec. 13-7). Many softcopy systems perform epipolar resampling after relative orientation so that the rows of pixels in both images lie along epipolar lines. This resampling can increase the efficiency of pattern matching even further.

References

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Boulianne, M., and C. Nolette: "Virtual Reality Applied to User Interfaces for Digital Photogrammetric Workstations," *Photogrammetric Engineering and Remote Sensing*, vol. 65, no. 3, 1999, p. 277.

Case, J. B.: ASP-DTM Symposium, *Photogrammetric Engineering and Remote Sensing*, vol. 44, no. 12, 1978, p. 1477.

Collins, S. H.: "Terrain Parameters Directly from a Digital Terrain Model," *Canadian Surveyor*, vol. 29, no. 5, 1975, p. 507.

Dowman, I. J.: "A Working Method for the Calibration of Plotting Instruments Using Computers," *Photogrammetric Record*, vol. 7, no. 42, 1973, p. 662.

——:"Model Deformation—An Interactive Demonstration," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 3, 1977, p. 303.

Graham, L. N., Jr., K. Ellison, Jr., and C. S. Riddell: "The Architecture of a Softcopy Photogrammetry System," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 8, 1997, p. 1013.

Helava, U. V.: "The Analytical Plotter—Its Future," *Photogrammetric Engineering and Remote Sensing*, vol.43, no. 11, 1977, p. 1361.

Kersten, T., and S. Haering: "Automatic Interior Orientation of Digital Aerial Images," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 8, 1997, p. 1007.

Khoshelham, K.: "Role of Tie Points in Integrated Sensor Orientation for Photogrammetric Map Compilation," *Photogrammetric Engineering and Remote Sensing*, vol. 75, no. 3, 2009, p. 305.
Mikhail, E. M.: "From the Kelsh to the Digital Photogrammetric Workstation, and Beyond," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 6, 1996, p. 680.
Mundy, S. A.: "Evaluation of Analytical Plotters for the Commercial Mapping Firm," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 4, 1984, p. 457.
Olaleye, J., and W. Faig: "Reducing the Registration Time for Photographs with Non-Intersecting Cross-Arm Fiducials on the Analytical Plotter," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 6, 1992, p. 857
Salsig, G.: "Calibrating Stereo Plotter Encoders," *Photogrammetric Engineering and Remote Sensing*, vol. 51, no. 10, 1985, p. 1635.
Zarzycki, J. M.: "An Integrated Digital Mapping System," *Canadian Surveyor*, vol. 32, no. 4, 1978, p. 443.
Zhang, Z., Y. Zhang, J. Zhang, and H. Zhang: "Photogrammetric Modeling of Linear Features with Generalized Point Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 74, no. 9, 2008, p. 1119.

Problems

12-1. List and briefly describe the four main categories of stereoplotters.

12-2. Describe the basic differences between stereoplotters with direct optical projection and instruments with mechanical or optical-mechanical projection.

12-3. Three basic orientation steps are necessary prior to using a stereoplotter. Name them and give the objective of each.

12-4. What is the focal length of the lens in a projector in the direct projection plotter of Fig. 12-2 whose principal distance and optimum projection distance are 152 and 760 mm, respectively?

12-5. Discuss the different viewing systems used in direct optical projection stereoplotters.

12-6. Outline the steps of interior orientation.

12-7. What are the advantages and disadvantages of using film diapositives in stereoplotters as opposed to glass diapositives?

12-8. Briefly describe the process of absolute orientation for a direct optical projection stereoplotter. Why must the steps be repeated?

12-9. Briefly describe the working components of an analytical stereoplotter.

12-10. Explain how the process of interior orientation for an analytical stereoplotter differs from that of an analog stereoplotter.

12-11. Repeat Prob. 12-10, except explain it for relative orientation.

12-12. Repeat Prob. 12-10, except explain it for absolute orientation.

12-13. Discuss the advantages of analytical plotters over analog plotters.

12-14. Compare softcopy plotters and analytical plotters, i.e., discuss their similarities and differences.

12-15. Describe the principle of the active polarizing screen method of stereoviewing for a softcopy plotter.

12-16. Describe the principle of the alternating LCS shutter display method of stereoviewing for a softcopy plotter.

12-17. Discuss the advantages of the fixed floating mark and moving images approach for stereoviewing over that of the moving floating mark and fixed images.

12-18. Briefly describe how automation can be applied to interior and relative orientation of a softcopy plotter.

CHAPTER 13

Topographic Mapping and Spatial Data Collection

13-1 Introduction

Mapping, and/or spatial data collection for GIS databases, can generally be categorized into either *planimetric* or *topographic* methods. Planimetric methods involve determining only the horizontal positions of features. The locations of these features are normally referenced in an *XY* coordinate system that is based upon some selected map projection (see Chap. 5). Topographic methods, on the other hand, include not only the location of planimetric details but also provide elevation information. Until recently, elevations have typically been represented by contours, which are lines connecting points of equal elevation. Now, however, elevations are often given in digital form; that is, *Z* coordinates are listed for a network of *X*, *Y* locations. This is particularly the case when data are being collected for GIS applications. This digital representation of elevations is called a *digital elevation model* (DEM), and this subject is discussed further in Sec. 13-6.

The purpose of the mapping, or spatial data collection, will normally dictate the level of detail required in representing features and the required level of accuracy as well. If the mapping is being done photogrammetrically by stereoscopic plotter, this in turn will dictate the flying height for the aerial photos, and hence also photo scale and map compilation scale. For example, maps used for the preliminary planning of a highway or transmission line would normally be compiled from high-altitude aerial photos which have a small scale. This would yield a small compilation scale, suitable for showing only general features and elevations with moderate accuracy. A map used for the design, construction, and operation of a utility, on the other hand, would generally be compiled from low-altitude photos, having a much larger scale. Thus compilation scale would be larger, and features and elevations would be shown in much greater detail and to a higher level of accuracy.

The concept of map scale and its relationship to level of detail that can be shown was well understood throughout the era when hard-copy maps were being compiled manually. That was the case because as map scale decreased, the number of features and contours that could be shown also had to decrease; otherwise, the map would have become congested to the point of being unreadable. Now, however, with the advent of digital mapping systems, the concept of mapping scale versus level of detail that can be shown is often not so obvious. With these systems, map products created from digital data can technically be plotted at any scale, and congestion can readily be removed by simply using the system's zoom feature. Nonetheless, the intended purpose of the mapping will still guide the level of detail presented, and this in turn will dictate photographic scale, compilation scale, and hence the positional accuracy of features shown.

13-2 Direct Compilation of Planimetric Features by Stereoplotter

In Chap. 12, fundamental operations of stereoscopic plotters were discussed. Regardless of the type of plotter employed, interior, relative, and absolute orientation (see Secs. 12-6 through 12-8) must be performed in order. Then compilation of the planimetric and/or topographic features within the "stereoscopic neat model" (see Sec. 18-7) can begin. Either of two basic methods may be employed in compiling topographic information with stereoplotters: (1) direct tracing from stereomodels to create hard-copy maps or (2) compiling digital files of topographic data from which computer-generated

maps are prepared. The first method is described in this section, while the second is covered in Secs. 13-4 through 13-6.

Prior to the existence of digital mapping systems, photogrammetric map compilation was accomplished exclusively by direct tracing of topographic features and contours from stereomodels. This process, although still occasionally performed, now has almost completely given way to digital mapping. In the direct tracing process, to locate a "point" feature such as a utility pole, the floating mark is placed on the object in the stereomodel, and the corresponding location of the tracing pencil marks the point's map location. "Linear" features such as roads and streams, and "area" features such as lakes and buildings, are drawn by placing the floating mark on the feature at some starting point, and then tracing the feature by moving the floating mark continuously along or around the object. The corresponding movement of the tracing pencil on the map sheet simultaneously locates the feature. While tracing features in this manner, it is absolutely essential that the floating mark be always kept in contact with the feature being traced; otherwise, its planimetric position will be incorrect. If a feature is not visible on one photo (or both), the floating mark cannot be properly set. Often when buildings are traced, the base of the building cannot be seen, and generally the roof edges will be traced. Slender vertical features such as utility poles can be plotted at either the top or the base, for if the floating mark is truly in contact with the pole, both will correspond to the same point.

As a general rule in direct compilation, planimetric details are traced first, followed by contouring. This is so because natural and cultural features have a very significant effect on the location and appearance of contours (e.g., the V's of contours crossing streams must peak in the stream and contours cross roads at right angles). An example of carelessly compiled contours or planimetry is shown in Fig. 13-1*a*, while its corresponding correct rendition is shown in Fig. 13-1*b*. Also in direct compilation, it is advisable to plot all features of a kind at once (e.g., all roads), before proceeding to another feature, as this reduces the likelihood of omissions.

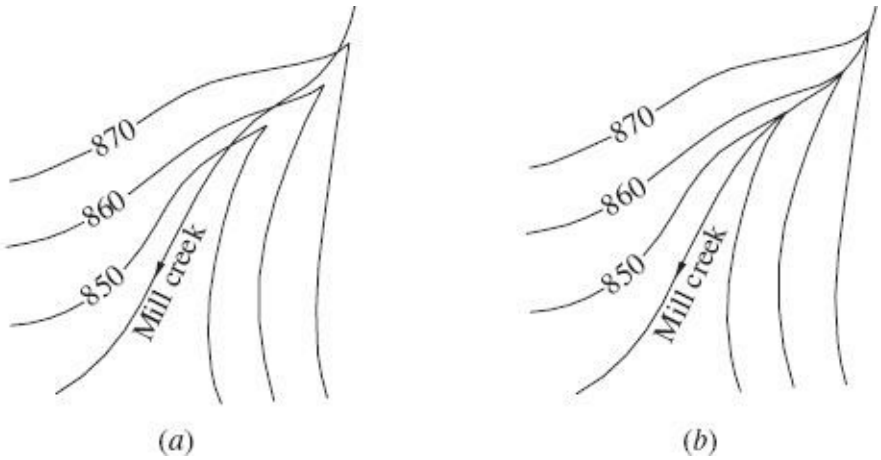


FIGURE 13-1 (*a*) Inconsistencies between planimetry and contour locations. (*b*) Consistent renditions of planimetry and contours.

13-3 Direct Compilation of Contours by Stereoplotter

Direct compilation of contours is a difficult operation that takes considerable skill and experience on the part of the operator. To trace contours, the floating mark is set at the elevation of a desired contour and is moved within the model until the mark coincides with the apparent terrain surface in the model. This will occur at any point where the desired contour elevation exists. The operator then moves the floating mark about the stereomodel as necessary to keep it constantly in contact with the terrain,

while the tracing pencil simultaneously follows the movement and records the contour. During the tracing of a contour, the operator must continuously look ahead to properly determine the direction in which to proceed.

In direct tracing of contours, it is recommended that all contours be completely compiled within a local area of the model, rather than attempt to trace a single contour across the entire model. To facilitate accurate contouring, the operator should become acquainted with the general shape of the terrain within the entire model before proceeding. This is usually most conveniently accomplished by viewing prints of the stereopair in three dimensions, using a stereoscope. Generally, contouring can be approached in much the same way as the assembly of a jigsaw puzzle. Easier areas and prominent features are compiled first, and the more difficult detail is filled in later. It can be helpful to initially locate and trace apparent drainage lines (ditches, swales, etc.) in order to guide the contouring operation. If necessary, these drainage lines can be subsequently erased during the map editing stage.

Some terrain areas such as flat expanses, regions of shadow, and areas of minimal image texture present particular difficulty when tracing contours. In these areas, it is sometimes necessary to determine “spot elevations” and interpolate contours from them. This is helpful because spot elevations can be read to significantly greater accuracy than direct contour tracing. In areas covered with trees or tall vegetation, it may be impossible to plot continuous contours. In these areas broken contours (contour lines plotted in open areas only) may be drawn; otherwise, spot elevations can be plotted and contours interpolated. Densely vegetated areas occasionally must be field-surveyed.

13-4 Digitizing Planimetric Features from Stereomodels

Features that are typically depicted in topographic mapping are many and varied. In general, however, in digital mapping any feature can be represented as consisting of: (1) a *point*, (2) a *line* (straight or curved), or (3) a *polygon* (closed shape). Horizontal and vertical control points, utility poles, and fire hydrants are examples of features that can be represented by points. Roads, railroads, streams, etc, can be shown as lines, while parcels of different land ownership or land-use types, large buildings, etc., can be represented by polygons.

As noted in Sec. 13-3, map compilation scale dictates the level of detail of the features to be included. This in turn may affect the manner of representing a particular feature. For example, on a 1:24,000-scale map, an 8 m-wide road would be plotted only 0.3 mm wide, and therefore it would be appropriately represented as a single line. On the other hand, the same road shown on a 1:600-scale map would be 13 mm wide, and therefore both edges of the road could be clearly shown without congestion.

In practice, features can be graphically represented by using different base-classes and sub-classes. This way, users can either select from predefined classes associated with specific features, or create their own. For example, a building sub-class is a type of rectangle base-class, and inherits its properties. That is, it can be represented by two perpendicular pairs of parallel lines and requires only three points to be defined. Thus, the user can simply select the building class to digitize rectangular buildings. In addition, the user may choose the rectangle base-class to digitize some other rectangular feature—such as an outdoor basketball court—that may not be a defined sub-class in the software. Further division is also possible beyond a two-class hierarchy. For instance, the rectangle class could be a sub-class of a polygon base-class. Predefined sub-classes typically have different graphical depictions depending on the features they represent in order to allow differentiation on the map. For example, while points can be used to define light poles, manholes, signs, and hydrants, each is assigned its own unique identifying symbol. While the classes described here refer to the shapes that are drawn during digitization, the features themselves may be grouped by category without regard to

the type of drawing class. The category only relates to the database representation of the feature–such as in GIS–while the class is used mainly for the digitizing process.

In digital mapping, code identifiers are used to keep track of the many different features that can be digitized from a stereomodel. Normally all features within one category are given a specific identifying code (number or letter). Then within that category, individual features may be further labeled with an additional identifier. Many different coding systems have been developed for categorizing map features based on type. These systems, while fundamentally similar, differ according to the primary purpose of the map product. Special considerations may also be required to fit the needs of individual projects. As an example, suppose an operator is digitizing all structures within a stereomodel, and in the mapping system being used, this category is identified by the number 3. Now within that category, buildings might be identified with the letter *A*, bridges with *B*, dams with *C*, retaining walls with *D*, etc. Thus, a bridge would be identified with a feature code of *3B*. The following feature categories (with typical individual features listed within each category) are presented only as a general guide. The list is not all-inclusive and may need to be supplemented on a project-specific basis.

1. *Nonimage data:* coordinate grid, property boundaries, horizontal and vertical control points
2. *Streets, highways, and related features:* edge of pavement, centerlines, curbs, medians, barrier walls, shoulders, guardrails, parking lots, alleys, driveways
3. *Other transportation:* railroad tracks, abandoned railroads, airport runways and taxiways, unpaved roads, trails, sidewalks, port facilities, locks
4. *Structures:* buildings, bridges, water towers, dams, fence lines, stadiums, retaining walls, antennas, concrete slabs, swimming pools
5. *General land use:* cemeteries, parks and recreation areas, agricultural areas, quarries, stockpiles, landfills
6. *Natural features:* lakes, ponds, rivers, streams, beaches, islands, wetlands, wooded areas, individual trees, bushes and shrubs, meadows
7. *Terrain elevation:* contours, spot elevations
8. *Drainage features:* ditches and swales, retention basins, culverts, headwalls, catch basins, curb inlets, storm sewer manholes, flared end sections, berms
9. *Utilities:* utility poles, power lines, telephone lines, transmission lines, substations, transformers, fire hydrants, gate-valve covers, sanitary manholes
10. *Signs and lights:* traffic signals, streetlights, billboards, street signs
11. *Project-specific:* varies

13-5 Representing Topographic Features in Digital Mapping

Traditionally most maps were traced directly from stereomodels and reproduced in hard copy by some type of printing process, while today the vast majority of maps are developed using computers and topographic data that have been compiled by digitizing from stereomodels. Many advantages accrue from the digital method of extracting topographic information. Perhaps the foremost advantage is that databases of geographic information systems require data in digital form; and in most cases, if proper formats and procedures are employed during the process of digitizing stereomodels, the resulting files can be entered directly without further modification into a GIS database. Other advantages are that digitized data afford greater flexibility of use. As examples, computerized mapping systems enable maps to be viewed on a screen, or they can be rapidly plotted in hard-copy form without the tedium of manual drafting. Furthermore, maps can be plotted using any desired scale and/or contour interval,

and the topographic data can be instantaneously transmitted electronically to remote locations. Because the information is collected according to feature category or "layer" in the digitizing process, individual layers can be plotted and analyzed separately, or overlaid with other layers for analyses. These are just a few of the advantages—there are many others.

The process of digitizing planimetric features from stereomodels is fundamentally the same, whether the operator is using a digitized *mechanical projection stereoplotter*, an *analytical plotter*, or a *softcopy plotter* (see Chap. 12). Of course, just as was the case of mapping by direct tracing from stereomodels, in digitizing data from stereomodels the instrument must be completely and accurately oriented prior to commencement of the digitizing process. To digitize an object within a stereomodel, the operator must bring the floating mark in contact with that object, enter the feature code, and then either push a button or depress a foot pedal. This causes the feature code and its *X*, *Y*, *Z* coordinates to be instantaneously stored in the computer. When digitizing linear or area features, a number of points along the line or polygon are digitized. This can be done very rapidly, since after the first point on any specific feature is identified and digitized, no further feature codes need be entered on that feature. On these types of features, most operators will pause momentarily when the button or foot pedal is depressed to ensure that the floating mark is on the desired object when the digitizing occurs. CAD systems associated with modern plotters provide for grouping of similar features into a specific layer in the drawing file. Generally, features are compiled in descending order of prominence. Point features that would appear smaller than about 1 mm at the intended map scale should be represented by appropriate symbols. Any notes or labels that the compiler feels are necessary to avoid later misidentification should also be made. Again, a set of paper prints and a stereoscope are essential to the compiler as an aid in identifying features.

13-6 Digital Elevation Models and Indirect Contouring

Digital representation of elevations in a region is commonly referred to as a *digital elevation model* (DEM). When the elevations refer to the earth's terrain, it is appropriately referred to as a *digital terrain model* (DTM). When considering elevations of surfaces at or above the terrain (tree crowns, rooftops, etc.), it can be referred to as a *digital surface model* (DSM). Any of these models can be represented as a regular grid or as a *triangulated irregular network* (TIN). While a DEM can technically be an elevation model of anything, in this book the term DEM will be used for elevation models of the earth's terrain in either regular grid or TIN form, unless otherwise specified.

In a regular grid DEM, spot elevations are determined for a uniformly spaced array of ground cells or groundels (See Sec. 9-3). The elevations (*Z* values) are digitally stored in a computer array. Figure 13-2 shows a three-dimensional representation of a rectangular DEM. In this figure, lines are drawn to connect the center points of the groundel array, which, by virtue of their varying *Z* values, appear as undulating terrain. Since the *z* values are stored as an array of numbers, they are compatible for computer processing using various algorithms. These algorithms include contour generation, volume and slope calculation, orthophoto generation (see Sec. 13-8), and depiction of three-dimensional views. Computer algorithms applied to rectangular arrays of elevations tend to be very straightforward in their implementation. They readily lend themselves to investigations such as frequency analysis and finite element analysis.

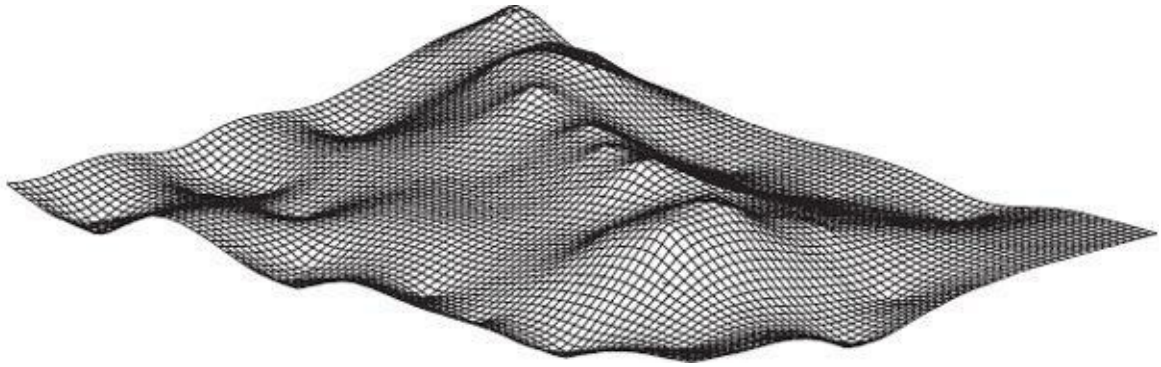


FIGURE 13-2 Three-dimensional representation of a rectangular DEM.

A disadvantage associated with regular grid DEMs is that for large areas of level or uniformly sloping terrain, a great deal of computer memory is wasted storing highly redundant elevation information. For example, if a groundel spacing of 1 m is used for a flat area of terrain 100 m^2, then 100 × 100 = 10,000 storage locations will be used to represent the elevations. Since the terrain is flat, all 10,000 values will be the same. More complex data structures such as *quadtrees* can be used in lieu of arrays to reduce storage requirements of regular grids; however, their implementation is far more complicated.

A TIN affords a more efficient representation of terrain in terms of data storage. With a TIN, spot elevations can be acquired at *critical points*, which are high points, low points, locations of slope changes, ridges, valleys, and the like. Once these spot elevations have been collected, together with their *X*, *Y* coordinates, lines are constructed between the points forming a system of triangles covering the surface. The triangles can be assumed to be planar facets which are sloped in three-dimensional space. Each triangle side has a uniform slope, and the system of triangles is called a TIN model. The method generally used to construct the TIN is known as a *Delaunay triangulation*. In a Delaunay triangulation, the lines are drawn between points in closest proximity to each other, without any lines intersecting. The resulting set of triangles has the property that for each triangle, the circle that passes through the three vertices contains no vertices of any other triangles.

A further generalization of a TIN allows the inclusion of *breaklines*. Breaklines are lines which have constant slope and are used where there are discontinuities in the terrain surface, such as streams, retaining walls, and pavement edges. Within the TIN, a breakline will form the sides of two adjacent triangles and will have no other lines crossing it. By using breaklines in a TIN, the terrain surface can be more faithfully represented, and contours generated from the TIN will give a more accurate portrayal of the surface features.

Figure 13-3 illustrates the concept of using a TIN as a DEM in order to determine contours. Figure 13-3*a* shows a contour map, which for purposes of illustration is assumed to accurately represent a particular terrain area. Note that the figure includes a hill near the left edge, a depression at the upper left, a steady rise to the upper right, and a stream (*AB*) which runs from the upper left to the lower right. Figure 13-3*b* shows a TIN (dashed lines) formed from a sample of 20 spot elevations (shown as black dots at all triangle vertices) along with the set of interpolated contours (solid lines). The contours are formed by interpolating along the sides of the triangles and connecting the interpolated points with straight lines. (Smooth curves can also be used to approximate the straight-line contours.) Although the figure gives a nominal depiction of the hill, depression, and rise, the contours associated with the stream (*AB*) are not properly represented. In the TIN shown in Fig. 13-3*c*, the original 20 spot elevations are still shown, but a breakline has been added to connect the endpoints (*A* and *B*) of the stream. Note that in this figure, none of the TIN lines cross the stream as they did in Fig. 13-3*b*. As such, the contours associated with the stream give a much better representation. An

even more accurate terrain representation is shown in Fig. 13-3*d*. In this figure, 72 spot elevations are used, in addition to the stream breakline. Note that the shapes and spacings of the contours in this figure are much closer to those in Fig. 13-3*a*, although some minor inconsistencies remain.

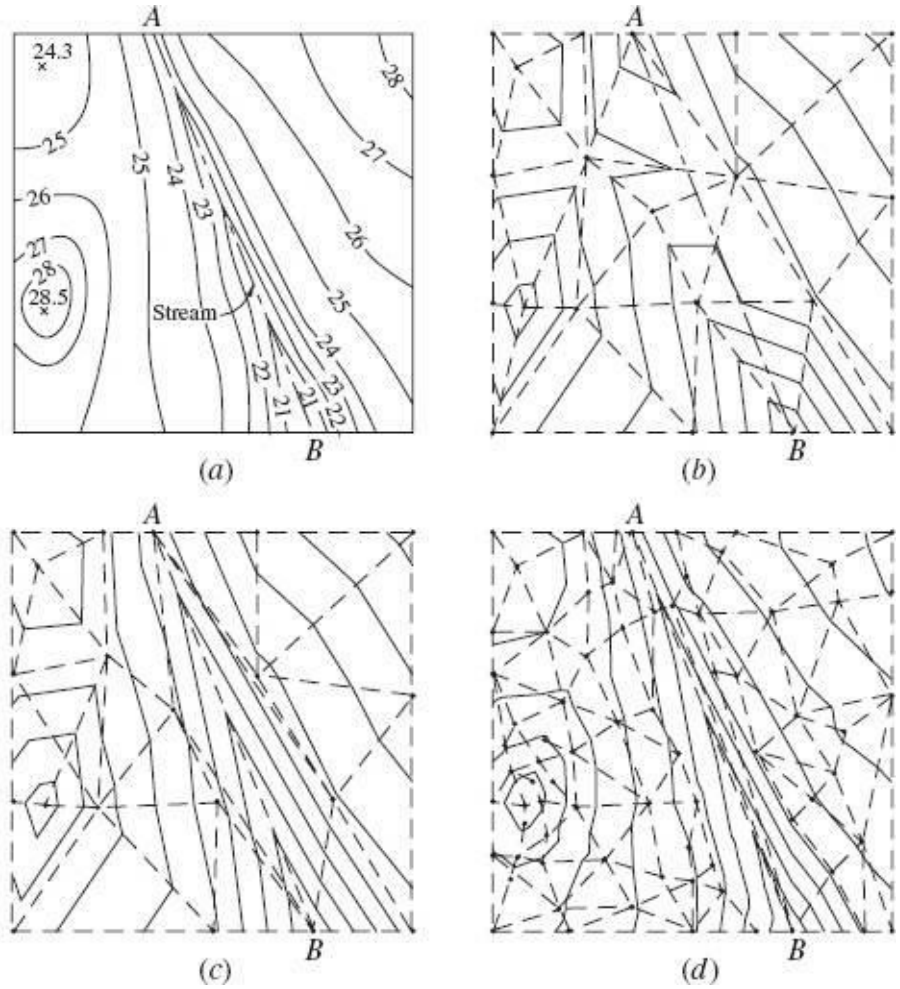


FIGURE 13-3 (*a*) Contour map showing accurate representation of terrain in an area. (*b*) Contours of the same area based on a TIN created from 20 data points, but without a breakline along stream *AB*. (Note the erroneous contour representations, especially in the area of the stream.) (*c*) Contours of the same area based on a TIN with 20 data points, but identifying stream line *AB* as a breakline. (Note the improvement in the contours in the area of the stream.) (*d*) Contours of the same area based on a TIN created from 72 data points, and with stream line *AB* identified as a breakline. [Note how well these contours agree with the accurate representation of (*a*).]

Because certain computational algorithms require the use of regular grids, it is sometimes necessary to convert a TIN to a regular grid DEM. This can be done by interpolating elevations between the TIN vertices at the centers of each grid cell. Various methods are available for performing this interpolation, including *nearest neighbor, inverse distance weighting, moving least squares, surface splines*, and *Kriging*. As its name implies, the nearest-neighbor method simply assigns the elevation of a grid cell to that of the nearest spot elevation in the TIN. In the inverse distance weighting approach, elevations at the grid cell locations are computed by a weighted average of spot elevations, where each weight is inversely proportional to the horizontal distance between the grid cell center and the spot elevation point. By this approach, spot elevations that are nearer to the grid cell will contribute more heavily to the weighted average. In the moving least squares approach, a local polynomial surface $z = f(x, y)$ is fitted to a small set of points nearest the grid cell being

interpolated. Once the polynomial coefficients have been computed by least squares, the z value (elevation) is computed at the given x, y of the grid cell. The procedure is then repeated for every additional grid cell, with different polynomial surfaces being computed for each. Surface splines are mathematically defined functions that generate a smoothly-varying surface that is forced to pass through each of the spot elevation points. Kriging is a statistical method which involves spatial correlation of points. The spatial correlation is specified a priori through the use of a *variogram* which quantifies the influence a spot elevation will have on points being interpolated. Proper definition of the variogram is important for generating an accurate interpolation. Figure 13-4 shows a three-dimensional view of a regular grid which has been interpolated by Kriging, using the spot elevations (and points along the breakline) of Fig. 13-3*d*.

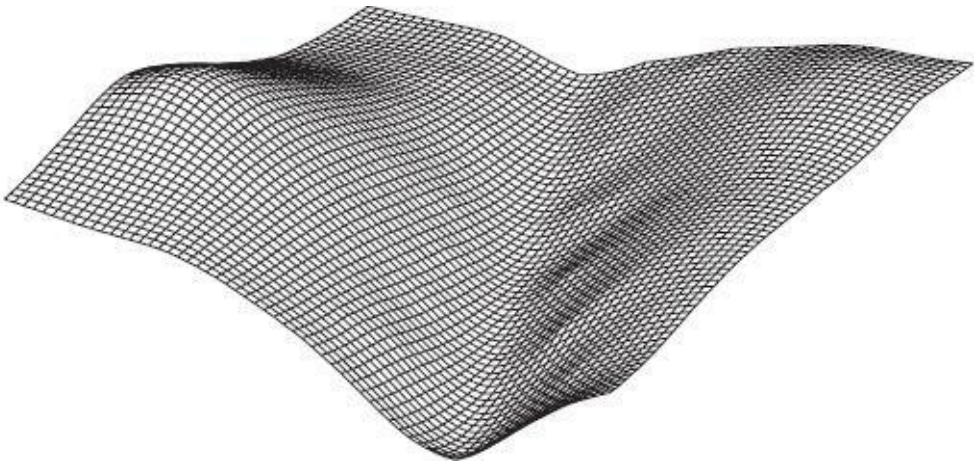


FIGURE 13-4 Three-dimensional view of a regular grid that has been interpolated from the data points of Fig. 13-3*d* by Kriging.

13-7 Automatic Production of Digital Elevation Models

An important application of softcopy stereoplotters, which relies upon digital image matching, is the automatic production of digital elevation models. In this process, a set of image points is selected from the left image and subsequently matched with corresponding points in the right image. The set of selected points is typically arranged in a nominal grid pattern, which results in a nearly uniform regular grid DEM. In general, the resulting DEM will differ somewhat from a perfectly regular grid due to tilt and relief displacements which exist in the images

Figure 13-5 is an illustration of the positions of DEM points within the overlap area of a stereopair of images. A boundary is chosen for the region in which the DEM points will be selected. The grid of selected points within the region on the left image is arranged at spacing of d_x and d_y in the x and y directions, respectively. Each of these points is matched with corresponding points from the right image, using image-matching techniques as discussed in the preceding section. Since the ground is generally not flat, the x parallax of DEM points will not be constant. As can be seen in Fig. 13-5, this varying x parallax causes the pattern of corresponding points from the right image to differ from a perfect grid arrangement.

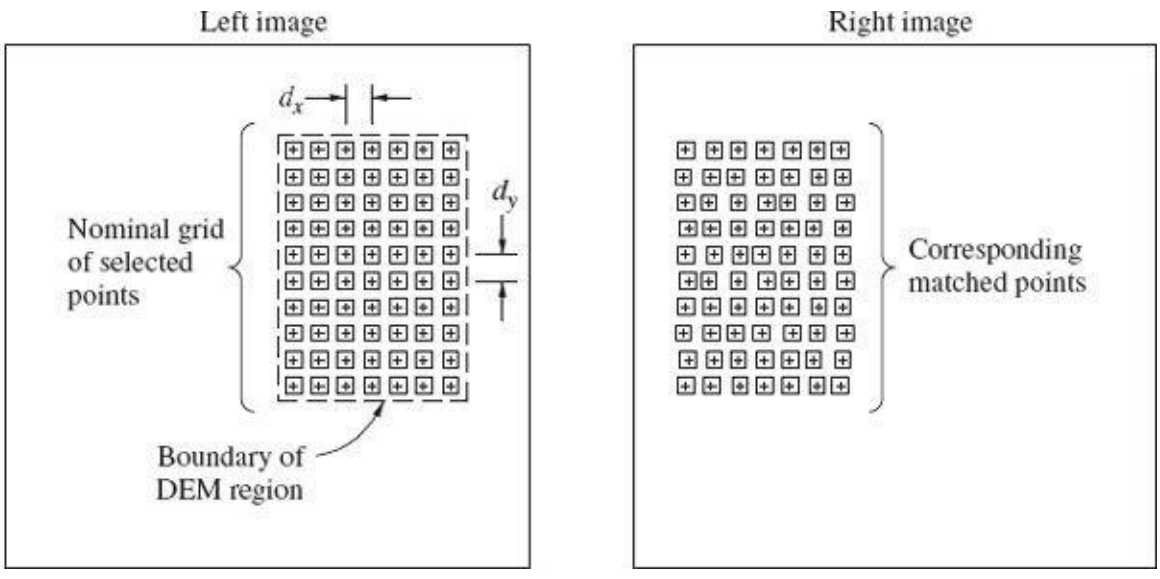


FIGURE 13-5 Positions of automatically matched DEM points in a stereopair.

Once the full set of DEM points has been matched and their object space coordinates calculated, their three-dimensional positions in the stereomodel can be represented by a set of floating marks superimposed upon the terrain. If desired, the plotter operator can then view these points in stereo, one at a time, and adjust the positions of points which do not appear to rest on the terrain. When the operator is satisfied with all vertical positions, the resulting DEM can be stored in a computer file. This DEM is then available for additional operations such as automatic contouring or the production of digital orthophotos, as discussed next.

13-8 Orthophoto Generation

An orthophoto is a photograph showing images of objects in their true orthographic positions. Orthophotos are therefore geometrically equivalent to conventional line and symbol planimetric maps which also show true orthographic positions of objects. The major difference between an orthophoto and a map is that an orthophoto is composed of images of features, whereas maps utilize lines and symbols plotted to scale to depict features. Because they are planimetrically correct, orthophotos can be used as maps for making direct measurements of distances, angles, positions, and areas without making corrections for image displacements. This, of course, cannot be done with perspective photos. As discussed in Chap. 20, orthophotos are widely used in connection with geographic information systems, where they serve as planimetric frames of reference for performing analyses, and are also used for generating layers of information for databases.

Orthophotos are produced from perspective photos (usually aerial photos) through a process called *differential rectification*, which eliminates image displacements due to photographic tilt and terrain relief. The process is essentially the same as standard rectification described in Secs. 10-14 and 10-15, except that it is performed independently at myriad individual, tiny surface patches or differential elements. In this way, rather than rectify the photograph to some average scale (which does not correct for relief displacements), each differential element is rectified to a common scale. Prior to the age of digital photogrammetry, complicated optical-mechanical devices were employed to produce orthophotos from film diapositives. Several of these are described in the second edition of this book. Modern orthophotos, however, are produced digitally, Softcopy systems are particularly well-suited for differential rectification. The essential inputs for the process of differential rectification are a DEM and a digital aerial photo having known exterior orientation parameters (ω, ϕ,

κ, X_L, Y_L, and Z_L). It is also necessary to obtain the digital image coordinates (row and column) of the fiducials so that a transformation can be computed to relate photo coordinates to digital image coordinates. A systematic application of the collinearity equations (see Sec. 11-4) is then performed to produce the orthophoto.

Figure 13-6 illustrates the collinearity condition for a particular groundel point *P* in the DEM. The *X* and *Y* coordinates of point *P* are based upon the row and column within the DEM array, and its *Z* coordinate is stored in the DEM groundel array at that position. Given the *X*, *Y*, and *Z* coordinates of point *P* and the known exterior orientation parameters for the photograph, the collinearity equations [see Eqs. (11-1) and (11-2)] can be solved to determine photo coordinates x_p and y_p. These photo coordinates define the position where the image of groundel point *P* will be found. Since photo coordinates are related to the fiducial axis system, a transformation must be performed on these coordinates to obtain row and column coordinates in the digital image. The transformed row and column coordinates will generally not be whole numbers, so resampling is done within the scanned photo to obtain the digital number associated with groundel point *P*.

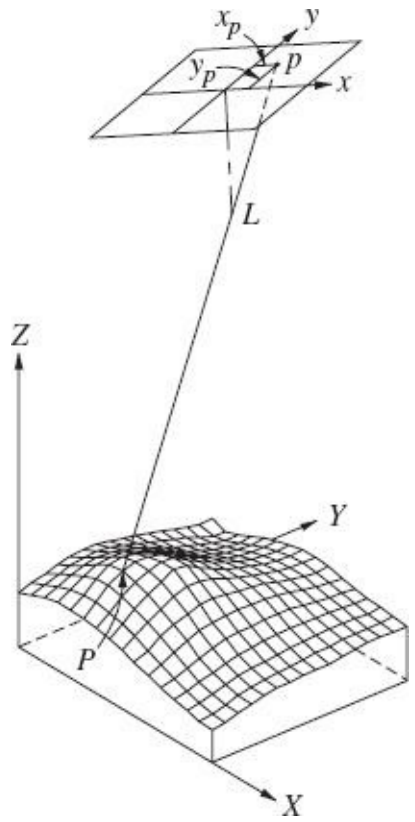


FIGURE 13-6 Collinearity relationship for a DEM point *P* and its corresponding image *p*.

The process of creating the digital orthophoto requires repetitive application of the collinearity equations for all the points in the DEM array. Figure 13-7 gives a schematic illustration of the process. In this figure, two arrays are shown in vertical alignment, where each of the groundels of the DEM array corresponds one-to-one with a pixel of the orthophoto array. The orthophoto array is initially empty, and will be populated with digital numbers (shown in the figure as x's) as the process is carried out. At each step of the process, the *X*, *Y*, *Z* coordinates of the center point of a particular groundel of the DEM are substituted into the collinearity equations as discussed in the preceding paragraph. The resulting photo coordinates are then transformed to row and column coordinates of the digital aerial photo. Resampling is performed to obtain a digital number, which is then placed into the corresponding pixel of the digital orthophoto. The process is complete when all pixels of the

orthophoto have been populated with digital numbers.

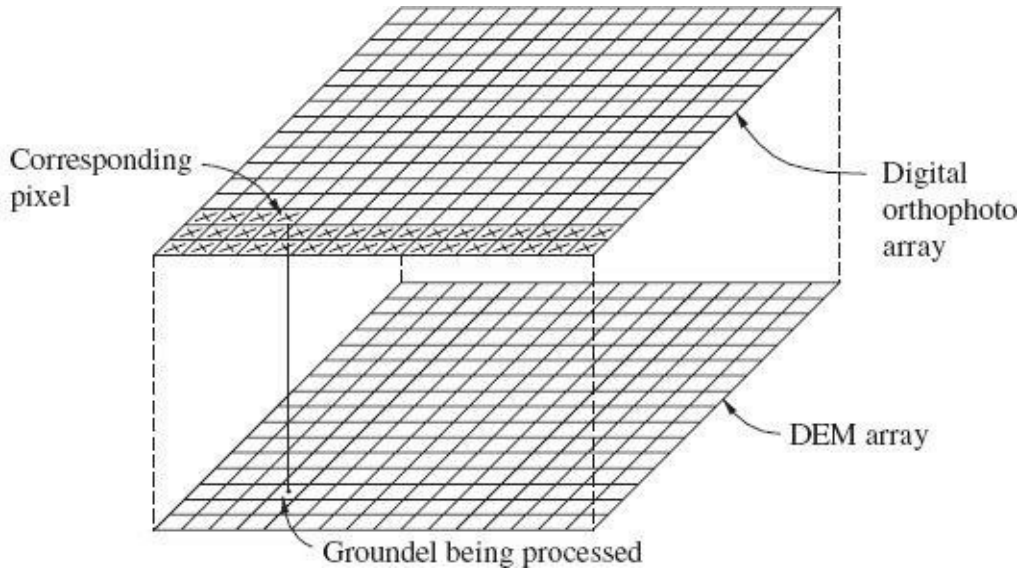
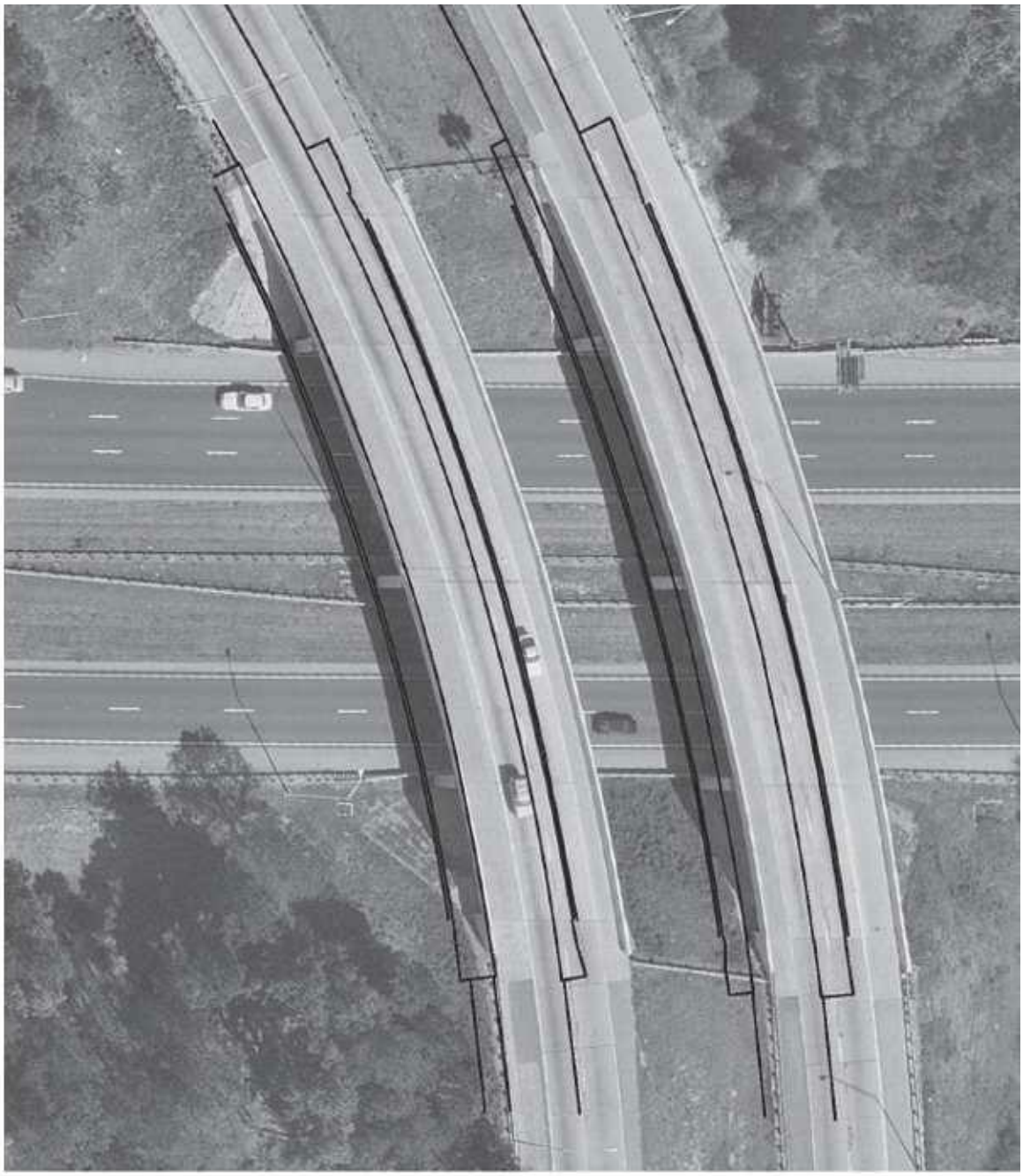


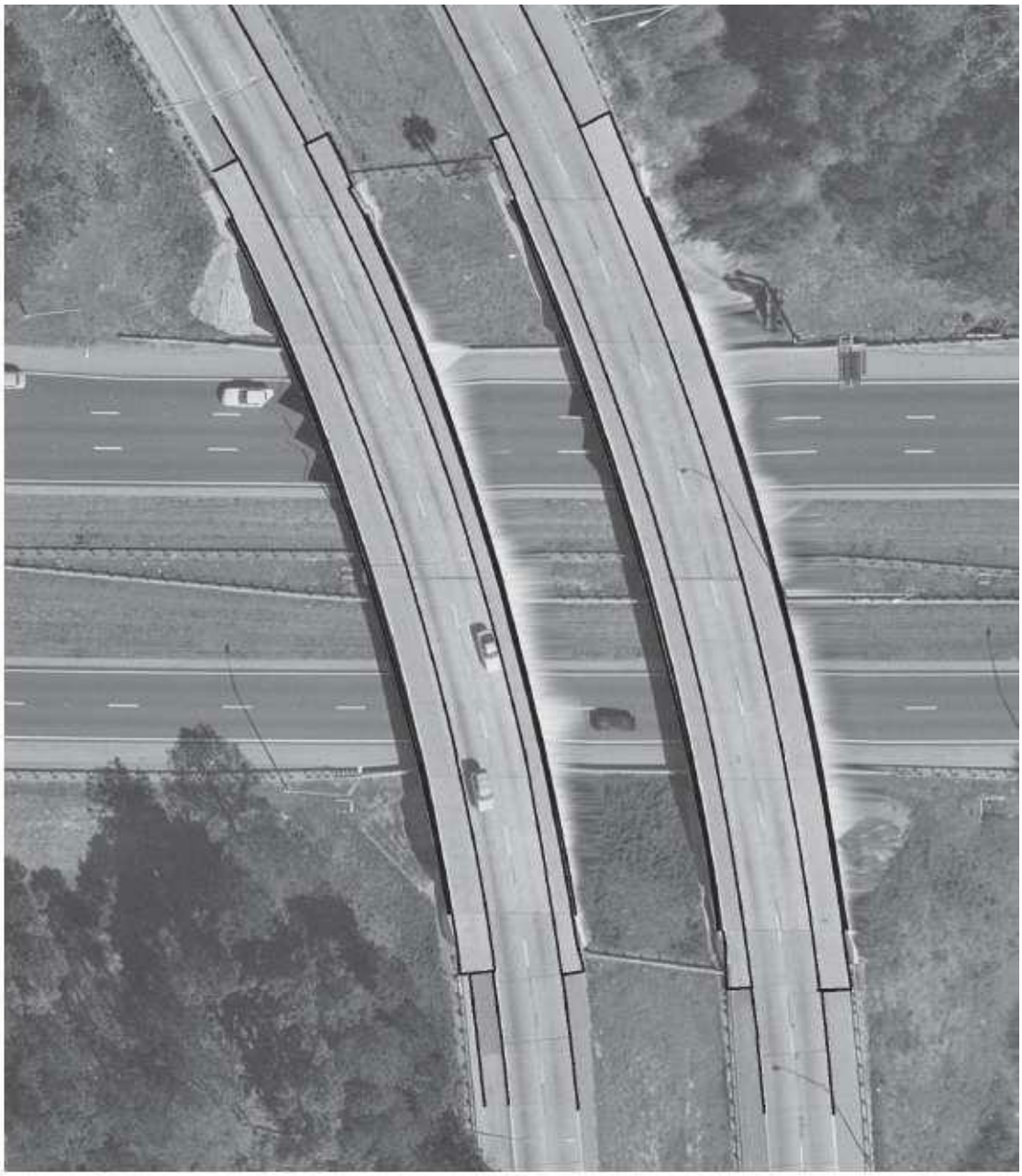
FIGURE 13-7 Schematic representation of digital orthophoto production process.

If the DEM used does not fully represent features on the ground surface, the resulting orthophoto can be dramatically distorted. Figure 13-8 shows the effect of DEM accuracy on the resulting orthophoto. The original photo, prior to orthorectification, is shown in Fig. 13-8*a*. Figure 13-8*b* shows the photo orthorectified with a DEM that does not include the bridge (i.e., elevation is the ground beneath the bridge), and Fig. 13-8*c* is the orthophoto resulting from a DEM that takes into account the height of the bridge above the ground.



(*a*)

(*b*)



(*c*)

FIGURE 13-8 Effect of DEM accuracy on orthorectification.

Orthophotomaps prepared from orthophotos offer significant advantages over aerial photos and line and symbol maps because they possess the advantages of both. On one hand, orthophotos have the pictorial qualities of air photos because the images of an unlimited number of ground objects can be recognized and identified. Furthermore, because of the planimetric correctness with which images are shown, measurements may be taken directly from orthophotos just as from line maps. When digital orthophotos are stored in a computer in their numeric form, they can be incorporated with a regular grid DEM having the same ground resolution to produce three-dimensional pictorial views. This is achieved by assigning the brightness values of the orthophoto pixels to corresponding cells of the DEM and manipulating the view angle. The result is a three-dimensional view of the terrain with the image draped over the top, as shown in Fig. 13-9.

FIGURE 13-9 Three-dimensional pictorial view of terrain obtained by combining a DEM and orthophoto.

13-9 Map Editing

After map manuscripts have been compiled with a stereoplotter, they are passed on to an editor who provides finishing touches. Since modern stereoplotter manuscripts are compiled in a CAD environment, the resulting digital files can be used directly by the editor at a computer workstation. The editor checks the manuscript for completeness and the use of proper symbols and labels. Endpoints of lines which delineate polygon features are forced to close properly, and buildings are squared up. Lines which continue between adjacent manuscript sheets are edge-matched so that they will be continuous. Proper line types (dashed, solid, etc.) are assigned along with appropriate line thickness. Depending upon project requirements and the features in the mapped area, additional edits may be performed. In a CAD environment, many of the editing functions can be achieved through automated processing, particularly edge-matching, closing polygons, and building squaring. After the manuscripts have been edited, the map should be checked to verify its compliance with the appropriate accuracy standards.

When the digital manuscript has been edited to its final form, the original stereoplotter manuscripts form a continuous and seamless composite map of the project area. The corresponding digital CAD file can be saved on a convenient data storage medium. When individual map sheets are required, the editor will extract areas of the digital manuscript in a "cookie-cutter" fashion and place them in an appropriate border and title block. A north arrow, scale bar, legend, notations, etc., are then placed on the map which can be subsequently plotted at the appropriate scale.

References

American Society of Photogrammetry: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

American Society for Photogrammetry and Remote Sensing: "Digital Elevation Model Technologies and Applications," *The DEM Users Manual*, 2d ed., Bethesda, MD, 2007.

Benjamin, S., and L. Gaydos: "Spatial Resolution Requirements for Automated Cartographic Road

Extraction," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 1, 1990, p. 93.
Carter, J. R.: "Digital Representations of Topographic Surfaces," *Photogrammetric Engineering and Remote Sensing*, vol. 54, no. 11, 1988, p. 1577.
Cowen, D. J.: "GIS versus CAD versus DBMS: What Are the Differences?" *Photogrammetric Engineering and Remote Sensing*, vol. 54, no. 11, 1988, p. 1551.
Doytsher, Y., and B. Shmutter: "A New Approach to Monitoring Data Collection in Photogrammetric Models," *Photogrammetric Engineering and Remote Sensing*, vol. 54, no. 6, 1988, p. 715.
Flood, M., and B. Gutelius: "Commercial Implications of Topographic Terrain Mapping Using Scanning Airborne Laser Radar," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 4, 1997, p. 327.
Fortune, S. J.: "A Sweepline Algorithm for Voronoi Diagrams," *Algorithmica*, vol. 2, 1987, p. 153.
Preparata, F. P., and M. I. Shamos, *Computational Geometry*, Springer-Verlag, New York, 1985.
Ramey, B. S.: "U.S. Geological Survey National Mapping Program: Digital Mapmaking Procedures for the 1990s," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 8, 1992, p. 1113.
U.S. Army Corps of Engineers: *Photogrammetric Mapping*, Available at: http://spatialdata.sam.usace.army.mil/organizations/survey/2002.
Wang, J., G. J. Robinson, and K. White: "A Fast Solution to Local Viewshed Computation Using Grid-Based Digital Elevation Models," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 10, 1996, p. 1157.

Problems

13-1. Explain the basic differences between planimetric mapping and topographic mapping.

13-2. Discuss the relationships between map scale and the level of detail that can be shown on a map.

13-3. Describe the process of using a stereoscopic plotter to create a hard-copy map by directly tracing planimetric features.

13-4. Describe the process of using a stereoscopic plotter to create a hard-copy map by directly tracing contours.

13-5. Why are feature codes necessary in digital mapping?

13-6. Describe the general process of digitizing planimetric features from stereomodels.

13-7. Cite several advantages of digital mapping over direct tracing of hard-copy maps.

13-8. Describe the following terms: (*a*) DEM and (*b*) TIN.

13-9. What are the slope characteristics that exist for the sides of all triangles of a TIN model?

13-10. What is Delaunay triangulation?

13-11. Describe breaklines, and discuss their importance in digital mapping.

13-12. How do orthophotos differ from normal aerial photos?

13-13. Discuss the similarities and differences between orthophotomaps and line and symbol maps.

13-14. What is the purpose of the map editing process? Name some specific conditions that are examined in the process.

CHAPTER 14
Laser Scanning Systems

14-1 Introduction

The laser was first manufactured in 1960 and some of its earliest applications included precise determination of the distance to faraway objects such as the lunar laser-ranging experiments, which were first carried out in 1962. However, it was not until the mid 1990s that advancements in GPS and inertial-aided navigation and georeferencing made large-scale mapping with lasers possible. Although the acronym Light Detection and Ranging (LiDAR) is widely used, the term *laser scanning* is more appropriate for describing most of the systems used for mapping. In addition to companies devoted solely to laser scanning, the complementary nature of laser scan data and aerial photography has led many companies who specialize mainly in photogrammetry to obtain laser scanning systems. Laser scanning has many advantages over photogrammetry, albeit with significant tradeoffs. For example, it is an active system, meaning that it provides its own energy, not relying on the sun to illuminate objects. This allows laser scanning to be performed at night and to avoid losing data due to sun shadows. It also has the advantage that it can sometimes penetrate tree canopy, allowing the ground to be mapped in forested areas. However, airborne laser scanning is at present less precise than airborne photogrammetric data with current systems yielding precisions in the neighborhood of 10-cm vertically and 15-cm horizontally under favorable conditions. It relies on georeferencing via airborne GPS and inertial systems, meaning that satellite configurations play a large role in how accurate the data is. In addition, laser scanning does not provide the spectral information that photography does, and the data sets can be unwieldy. These issues have led to the common practice of simultaneously acquiring both types of sensor data and fusing them to create combined data sets. Figure 14-1 shows aerial photography “draped” over laser scan data.

FIGURE 14-1 Combined aerial photography and laser scan data. (*Image courtesy of Optech*)

14-2 Principles and Hardware

The low divergence of laser beams allows one to aim light precisely at small areas. The main principle behind laser mapping systems is that a laser pulse is transmitted to an object or terrain and its return signal is detected a split second later. By accurately measuring the time delay between the transmission and return, and applying the speed of light, the distance from the scanner to the terrain or object point can be determined as shown in Eq. (14-1). In Eq. (14-1), t is the time between when the pulse is emitted and when the return is detected, and c is the speed of light (in the atmosphere). Therefore, the basic sensor hardware for laser mapping systems is composed of a laser emitter, a laser receiver, and a timing unit.

$$D = c\frac{t}{2} \tag{14-1}$$

In order to cover large areas with multiple laser pulses, the direction of the laser must be changed rapidly and precisely. Different scanning mechanisms are used to achieve this including oscillating mirrors and rotating multifaceted mirrors. Figure 14-2 shows the point pattern from an airborne oscillating mirror scanning system (left), and that from a rotating multifaceted mirror scanning system (right). Note the zigzagging pattern generated by the oscillating mirror compared to the parallel lines from the rotating mirror. The oscillating mirror system leads to smaller point spacing in the across-flight direction, and larger point spacing in the in-flight direction at the edge of the scan due to the mirror slowing down to change directions. Some systems modify the pulse rate during acquisition to compensate for this effect.

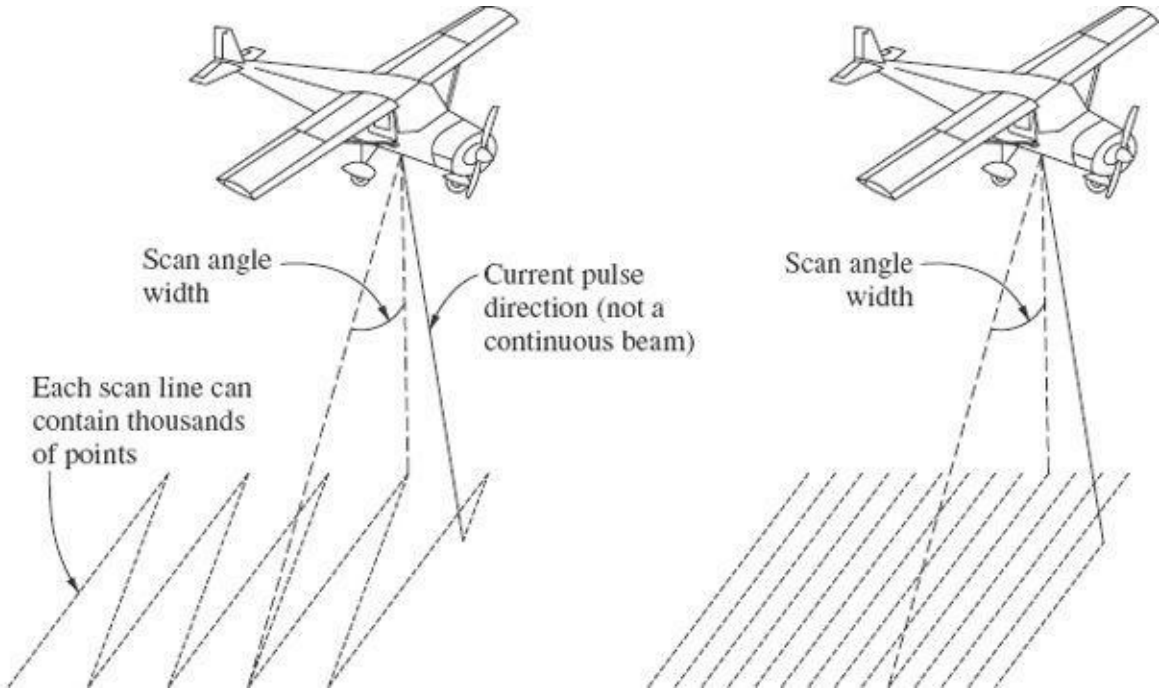


FIGURE 14-2 Scan pattern from an oscillating mirror laser scanning system.

Example 14-1
A laser mounted on an airplane emits a pulse that reflects off a target and returns to the sensor in .0066000 millisecond. Assume that the pulse is aimed directly down to the datum surface, that the speed of light is 299, 792, 458 m/s, and that the laser is at exactly 1 km above the datum. What is the object's height above datum?

Solution

1. By Eq.(14-1),

$$D = (299{,}792{,}458 \text{ m/s})\frac{(0.0066000 \text{ ms}) \times (.001 \text{ s/ms})}{2}$$

$$= 989.32 \text{ m}$$

2. Subtract the distance D from the distance to the datum.

$$1{,}000 \text{ m} - 989.32 \text{ m} = 10.68 \text{ m} \quad \blacktriangle$$

Example 14.2
Suppose that in Example 14-1 there is a 5 nanosecond (0.000005 millisecond) error in the timing device. What will be the error in the calculated object height above datum?

Solution By Eq. (14-1),

$$D = (299{,}792{,}458 \text{ m/s})\frac{(0.000005 \text{ ms}) \times (.001 \text{ s/ms})}{2}$$

$$= 0.75 \text{ m} \quad \blacktriangle$$

14-3 Airborne Laser Scanning

The most common type of laser scanning is *airborne laser scanning* (ALS). In addition to the laser scanning mechanism, ALS systems are composed of georeferencing hardware including an inertial measurement device and a GPS receiver. As the aircraft travels over the survey area scanning the terrain below, the inertial navigation device periodically records the attitude angles (roll, pitch, and yaw; or omega, phi, and kappa). In addition, the interfaced GPS receiver is periodically recording the X, Y, Z positions of the antenna. After the signals from the three devices are processed and combined, the laser pulses essentially define vector displacements (distances and directions) from specific points in the air, to points on the ground. Equation (14-2) shows relationship between point position on the ground and the orientation and angle of the laser scanner, where: X_A, Y_A, and Z_A are the point coordinates; X_S, Y_S, and Z_S are the scanner position coordinates; M is the rotation matrix from the scanner to the ground system; and α is the pointing angle. Since the laser device can generate pulses at a very fast rate (thousands of pulses per second), the result is a dense pattern of measured X_A, Y_A Z_A points on the terrain. Figure 14-3 shows an ALS system with scanning and georeferencing hardware.

$$\begin{pmatrix} X_A \\ Y_A \\ Z_A \end{pmatrix} = \begin{pmatrix} X_S \\ Y_S \\ Z_S \end{pmatrix} + M \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) \\ 0 & -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -D \end{pmatrix} \quad (14\text{-}2)$$

FIGURE 14-3 Optech Orion airborne laser scanning system. (I*mage Courtesy of Optech*)

Similar to photogrammetric flight plans, ALS flights are flown so that overlapping scanned *swaths* are produced for the survey area. This is to ensure full coverage of the area, avoiding data gaps caused by terrain relief and to allow for calibration. Figure 14-4 shows an ALS swath pattern illustrating the impact of terrain change on swath width. Note that swath width is a function of the height of the scanner above terrain and the *scan angle*, the size of the angle swept by the laser. In addition to overlapping parallel swaths, a common procedure is to fly a few swath lines perpendicular to the other lines for use in the calibration of the scanning mirror and other hardware.

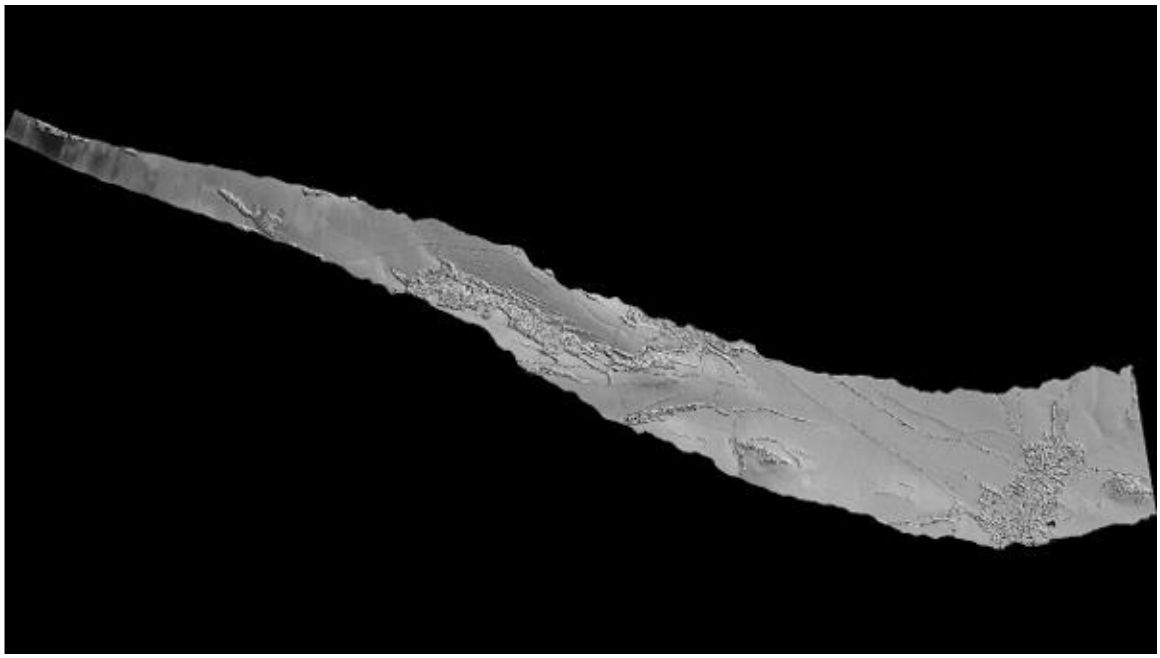


FIGURE 14-4 Swath pattern from an ALS system over variable terrain. (*Courtesy of RIEGL USA, Inc.; http://www.rieglusa.com*.)

Example 14-3

An airplane carrying an ALS system emits a laser pulse with a pointing angle of α = 5.000° that takes .0051110 millisecond to reach an object on the ground and return to the sensor. At the same time, an onboard GPS-INS system measures the position of the laser coordinates as X = 100.00 m, Y = 100.00 m, Z = 1000.00 m, and the orientation as ω = ϕ = κ = 0. What is the location of the object on the ground?

Solution

1. Solve for D using Eq. (14-1).

$$D = (299{,}792{,}458 \text{ m/s})\frac{(.0051110 \text{ ms}) \times (.001 \text{ s/ms})}{2}$$

$$= 766.12 \text{ m}$$

2. Solve for X_A, X_A, Z_A using Eq. (14-2). Note that M is the identity matrix because all rotations are zero and can be removed from the equation.

$$\begin{pmatrix} X_A \\ Y_A \\ Z_A \end{pmatrix} = \begin{pmatrix} 100.00 \text{ m} \\ 100.00 \text{ m} \\ 1000.00 \text{ m} \end{pmatrix}$$

$$+ M \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(5.000°) & \sin(5.000°) \\ 0 & -\sin(5.000°) & \cos(5.000°) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -766.12 \text{ m} \end{pmatrix}$$

$$= \begin{pmatrix} 100.00 \text{ m} \\ 100.00 \text{ m} \\ 1000.00 \text{ m} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.83867 & 0.54464 \\ 0 & -0.54464 & 0.83867 \end{pmatrix}$$

$$\times \begin{pmatrix} 0 \\ 0 \\ -766.12 \text{ m} \end{pmatrix} = \begin{pmatrix} 100.00 \text{ m} \\ -33.23 \text{ m} \\ 236.80 \text{ m} \end{pmatrix} \blacktriangle$$

Example 14-4
An ALS system has a pulse rate of 20 kHz, a scan rate (left to right and back) of 50 Hz, a flying height of 1000 m, and a scan angle of 40°. What is the estimated spacing of pulses along the ground?

Solution

1. Calculate the number of pulses per line.

$$pulses\ per\ line = \frac{(20 \text{ kHz}) \times (1000 \text{ Hz/kHz})}{(50 \text{ Hz}) \times 2}$$

$$= 200$$

2. Calculate W, the swath width on the ground.

$$W = 2 \times (1000 \text{ m}) \times \tan(40°/2) = 728 \text{ m}$$

3. Calculate pulse spacing.

$$\begin{aligned} \textit{pulse spacing} &= W / \textit{pulses per line} \\ &= 728 \text{ m} / 200 = 3.6 \text{ m} \end{aligned}$$ ▲

14-4 Terrestrial Laser Scanning

Terrestrial laser scanning (TLS) is usually done using a static platform. That is, as opposed to ALS, the scanner is set up at a location and scans the survey area from this stationary position. Since there is no movement of the platform, either the scanner's head must rotate or an additional mirror must be used to scan. Figure 14-5 shows a terrestrial laser scanner.

FIGURE 14-5 RIEGL VZ400 terrestrial laser scanner. (*Courtesy of RIEGL USA, Inc.; http://www.rieglusa.com.*)

TLS data is typically georeferenced by scanning targets with known datum coordinates. The 3D coordinate transformation (see section C-7) from the scanner system to the datum system is calculated using the target points, and is then applied to all other points in the scan. Similarly, scans from different positions can be combined using the same process with common targets between scans. TLS data may also be georeferenced by fixing a GPS antenna to the scanner itself. In this case, multiple scans are combined using the method described above, and the positions of the antenna in the combined scanner system can be used to solve for the transformation from the combined scanner system to the datum system. Note that you must have a minimum of three scanner setups to use this method. Another popular method for combining scans is by using surface matching. The set of points for one area from one scan is compared with the points for the same area in another scan, and the relative orientation between scans is calculated based on the criteria of how well the scans "fit" together. A common method for surface matching is the iterative closest point (ICP) algorithm, in which the relative orientation of two scans is iteratively updated towards a solution criterion that the sum of the distance between paired points from different scans is minimized.

FIGURE 14-6 RIEGL VMX450 mobile laser mapping system. (*Courtesy of RIEGL USA, Inc.; http://www.rieglusa.com*.)

Terrestrial scanning may also be collected using a mobile platform. In this case, similar to aerial mapping, an onboard GPS receiver and inertial system must be used. Mobile TLS combines the ability of static TLS to collect 3D data from a ground perspective with the ability of ALS to collect data over large areas in a short amount of time. Since it uses direct georeferencing, it avoids the need for control observations—a major limitation of its static counterpart. However, the data from mobile TLS is generally much less accurate than static, and results are highly dependent on the visibility of satellites. Mobile TLS is mostly used for mapping urban streets and highways; however, as more sophisticated processing is developed it is expected to be used in a wider range of applications.

14-5 Laser Scan Data

Laser scan data consists of sets of 3D coordinates called *point clouds*. Because of the high rate of collection, the size of these point clouds can contain millions of coordinates. Since the laser pulse can penetrate through certain types of vegetation and other objects, each pulse may yield multiple returns. For instance, when a pulse is emitted, a portion of it may hit leaves at the top of a tree and return to the sensor, the remaining pulse may strike a limb on the same tree and return to the sensor, and the rest of the pulse may hit the ground and reflect back to the sensor. These multiple returns are each recorded. Figure 14-7 illustrates multiple returns from a single pulse.

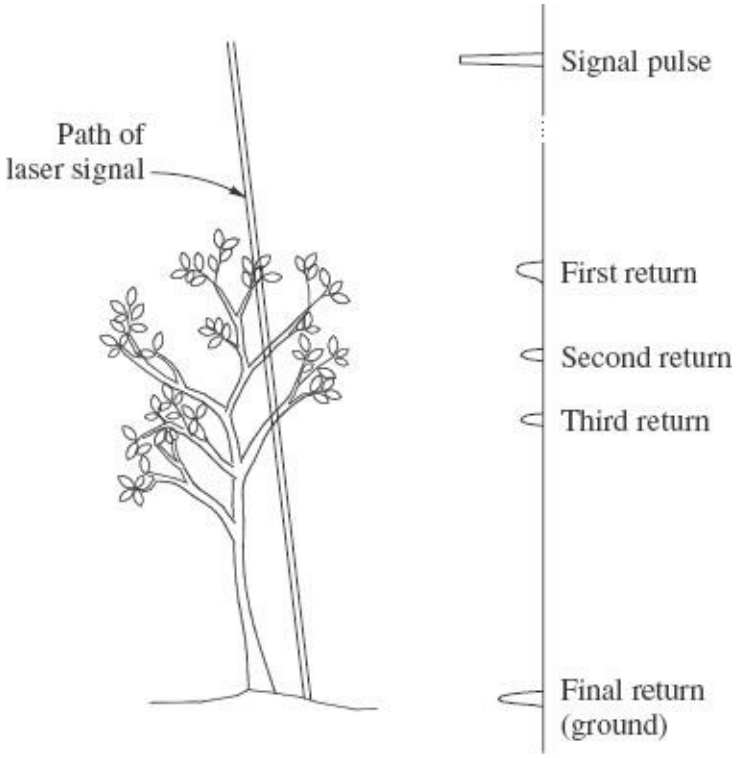
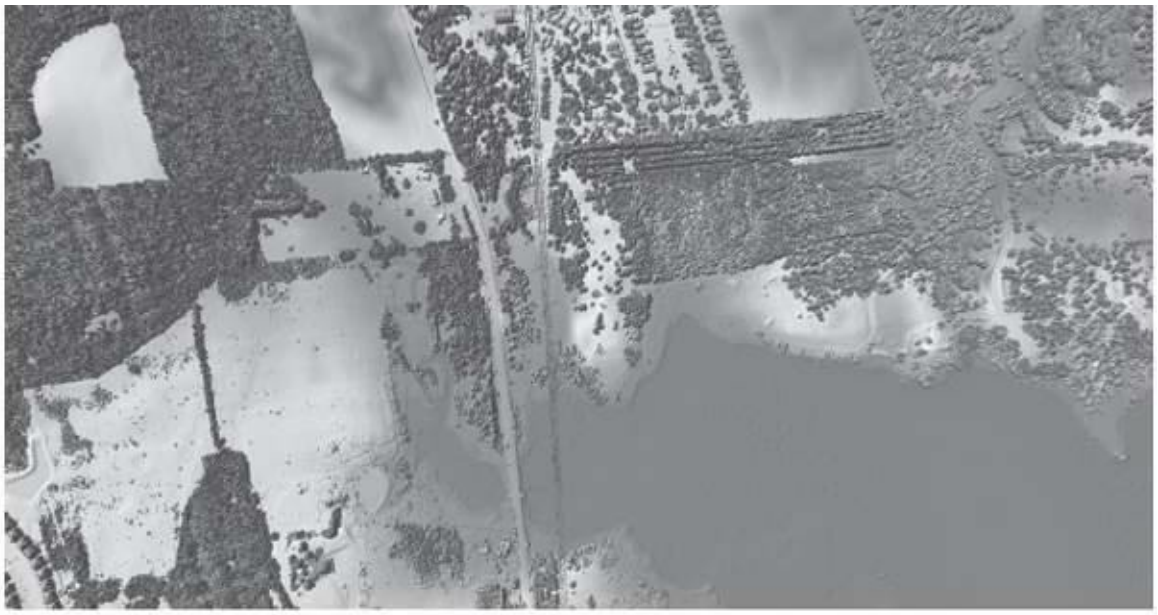
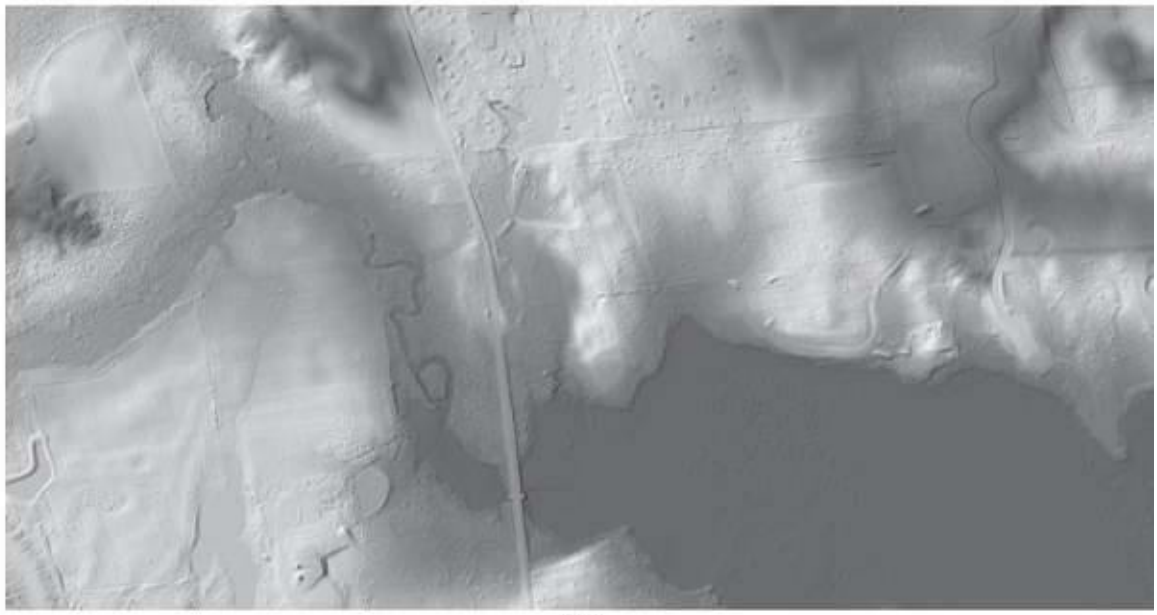


FIGURE 14-7 Multiple returns from a single laser pulse.

After the initial processing of the point cloud to determine the real-world, datum coordinates of the points, operations may be performed on the data to provide specific information sought by the user. For instances, a common procedure is to filter the data so that only last returns are used. This results in a bare-earth DEM, an exceptional example of the advantage of laser scanning to penetrate canopy. Figure 14-8*a* shows before and *b* shows after images of a filtered point cloud. The data is also commonly interpolated to create a gridded DEM. This DEM can be used to create contour maps and other topographic products.



(*a*)



(*b*)

FIGURE 14-8 Unfiltered and filtered ALS data. (*Image Courtesy of Optech.*)

Laser scanning systems can be classified as either *discrete* or *full-waveform* with respect to how returns are recorded. Discrete sensors record the ranges for each return, while full waveform sensors record a quantized waveform of the returning echo. Using full waveform sensors allows for the extraction of extra information from the object being scanned like surface roughness, and can facilitate the filtering of terrain with short vegetation. The major downside to full-waveform systems is the immense amount of data that is recorded. Similarly, the intensities, or the strengths of the returning signals, are usually recorded and provide even further information about the scanned object and can be used to aid in classification of the data.

14-5 Error Evaluation

The accuracy of the position of 3D points collected by ALS is most affected by the precision of the processed GPS data. As better kinematic GPS algorithms emerge, accuracy is expected to increase. Although a nominal accuracy is associated with ALS data, the discrimination of the height of objects

that are close together can be problematic due to the footprint, or laser beam diameter on the ground. In addition to the footprint of the laser, the pulse length can contribute to ambiguity in ALS data. This is particularly an issue with getting bare earth data from areas with low vegetation. For example, if a pulse is 3 m long, and there is vegetation on the ground that is 1.5 m high, there will be an overlap in the returns from the vegetation and the ground. The convolution of these returns can cause a bias in the determination of the ground (second return) height. The use of full waveform sensors can help alleviate this problem.

Similar to airborne photogrammetry using direct georeferencing, boresight and leverarm calibrations must be found to get accurate results in ALS, and contribute significantly to errors in the data if they are coarsely resolved. There may also be a scaling error of the angle traversed by the scanning mechanism. In other words, there may be errors stemming from the device that records the pointing angle of the laser beam.

References

Ackermann, F.: "Airborne Laser Scanning—Present Status and Future Expectations", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, 1999, p. 64.

Baltsavias, E. P.: "Airborne Laser Scanning: Basic Relations and Formulas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, 1999, p. 199.

Barber, D., J. Mills, and S. Smith-Voysey: "Geometric Validation of a Ground-based Mobile Laser Scanning System," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 63, 2008, p. 128.

Hodgson, M. E., and P. Bresnahan: "Accuracy of Airborne Lidar-Derived Elevation: Empirical Assessment and Error Budget," *Photogrammetric Engineering and Remote Sensing*, vol. 70, no. 3, 2004, p. 331.

Pfeifer, N., and C. Briese: "Laser Scanning—Principles and Applications," Proceedings of 3rd International Exhibition & Scientific Congress on Geodesy, Mapping, Geology, Geophisics, Cadaster, Novosibirsk, Russia, 2007, p. 311.

Shrestha, R., W.E. Carter, M. Lee, P. Finer, and M. Sartori: "Airborne Laser Swath Mapping: Accuracy Assessment for Surveying and Mapping Applications," *Surveying and Land Information Systems*, vol. 59, no. 2, 1999, p. 83.

Wehr, A., and Uwe Lohr: "Airborne Laser Scanning—an Introduction and Overview," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, 1999, p. 68.

Shan, J., and A. Sampath: "Urban DEM Generation from Raw Lidar Data: A Labeling Algorithm and its Performance," *Photogrammetric Engineering and Remote Sensing*, vol. 71, no. 2, 2005, p. 217.

Problems

14-1. Discuss the differences between laser scanning and photography and explain the benefits of collecting both at the same time.

14-2. What are some unique properties of lasers that make them useful for measuring distances?

14-3. Compare oscillating mirror to rotating multifaceted mirror scanning devices by listing advantages and disadvantages of each.

14-4. A laser mounted on an airplane emits a pulse that reflects off a target and returns to the sensor

in 0.0062003 millisecond. Assume that the pulse is aimed directly down to the datum surface, that the speed of light is 299, 792, 458 m/s, and that the laser is at exactly 1.2 km above the datum. What is the object's height above datum?

14-5. Repeat Prob. 14-4, except the return time is 0.00054923 millisecond, and the laser is at 1.5 km above datum.

14-6. An airplane carrying an ALS system emits a laser pulse with a pointing angle of $\alpha = 14.723°$ that takes 0.0051245 millisecond to reach an object on the ground and return to the sensor. At the same time, an onboard GPS-INS system measures the position of the laser coordinates as X = 100.00 m, Y = 100.00 m, and Z = 1000.00 m, and the orientation as $\omega = \Phi = \kappa = 0°$. What is the location of the object on the ground?

14-7. Repeat Prob. 14-6, except the orientation angles are $\omega = 0.51224°$, $\phi = -0.05992°$, $\kappa = 178.6729°$.

14-8. Discuss the assumptions that are made about the relative location of the laser emitter and laser sensor in Probs. 14-4 through 14-7.

14-9. An ALS system has a pulse rate of 33 kHz, a scan rate (left to right and back) of 70 Hz, a flying height of 600 m, and a scan angle of 40°. What is the estimated spacing of pulses along the ground?

14-10. Repeat Prob. 14-9, except the pulse rate is 167 kHz, the scan rate is 100 Hz, the flying height is 1100 m, and the scan angle is 50°.

CHAPTER 15

Fundamental Principles of Digital Image Processing

15-1 Introduction

Digital image processing in general involves the use of computers for manipulating digital images in order to improve their quality and/or modify their appearance. In digital image processing, the digital number (see Sec. 2-12) of each pixel in an original image is input to a computer, with its inherent row and column location. The computer operates on the digital number according to some preselected mathematical function or functions, and then stores the results in another array which represents the new or modified image. When all pixels of the original image have been processed in this manner and stored in the new array, the result is a new digital image.

Many different types of digital image processes can be performed. One type falls under the general heading of *preprocessing operations*. These are generally aimed at correcting for distortions in the images which stem from the image acquisition process, and they include corrections for such conditions as scanner or camera imperfections and atmospheric refraction. These procedures are discussed in Chap. 4. Another type of digital image processing, called *image enhancement*, has as its goal, the improvement of the visual quality of images. Image enhancement makes interpretation and analysis of images easier, faster, and more accurate; and thus it can significantly improve the quality of photogrammetric products developed from digital images, and reduce the cost of producing them. Digital orthophotos in particular benefit significantly from the improved image quality that results from image enhancements.

A third type of digital image processing, called *image classification*, attempts to replace manual human visual analysis with automated procedures for recognizing and identifying objects and features in a scene. Image classification processes have been widely used in a host of different interpretation and analysis applications, as well as in the production of a variety of thematic maps. They are also used in automated soft copy mapping systems (see Chap. 12). A final type of digital image processing, *data merging*, combines image data for a certain geographic area with other geographically referenced information in the same area. The procedures may overlay multiple images of the same area taken at different dates—a technique which is very useful in identifying changes over time, such as monitoring a forest fire or following the spread of a disease in a certain tree species. The procedures can also combine image data with nonimage data such as DEMs, land cover, and soils. These types of digital image processing are extremely important in the operation of geographic information systems.

In this chapter, some concepts that are fundamental to digital image processing are presented, and basic procedures for a few of the more common types of digital image processes are introduced. Examples are given to illustrate their effects.

15-2 The Digital Image Model

To comprehend the subject of digital image processing, an understanding of the concept of the digital image model is useful. In a mathematical sense, a digital image model can be considered as a mathematical expression which yields image density as a function of x, y (row, column) position. This is analogous to a regular grid DEM which yields elevation of a terrain surface as a function of X, Y position.

Digital images are imperfect renditions of the object scenes they portray. Their imperfections stem from a host of sources, including the imaging system, signal noise, atmospheric scatter, and shadows. The primary degradation of an image is a combined systematic blurring effect resulting from aberrations of the lens, resolution of the recording medium (e.g., CCD Array/film), and, to some extent, atmospheric scatter. The effect of these combined factors can be specified in terms of the *point-spread function*. A point-spread function can be thought of as the blurred image that would result from a perfect point source of light such as a star. This point spread function can be represented as a mathematical expression which models the systematic image imperfection, and which was applied to the ideal image through a process known as *convolution*. (The subject of convolution in image processing is discussed in Sec. 15-5, and examples are given.)

The cumulative effect of the systematic image degradations can be represented in a general sense by

$$I(x, y) = O(x, y) * P(x, y) + N(x, y) \tag{15-1}$$

In Eq. (15-1), $I(x, y)$ is the actual image model as represented by the digital image, $O(x, y)$ is the theoretical ideal image model of the object, $P(x, y)$ is the point-spread function, $N(x, y)$ is signal noise, and * (asterisk) indicates the convolution operation. Figure 15-1 illustrates convolution of the ideal image model by the point spread function. Notice how the high frequency features in the original object are are smoothed in the output image. Spatial frequency is explained in the next section. One goal of image processing is to negate the effects of image noise and the point-spread function to recover the theoretical ideal image.

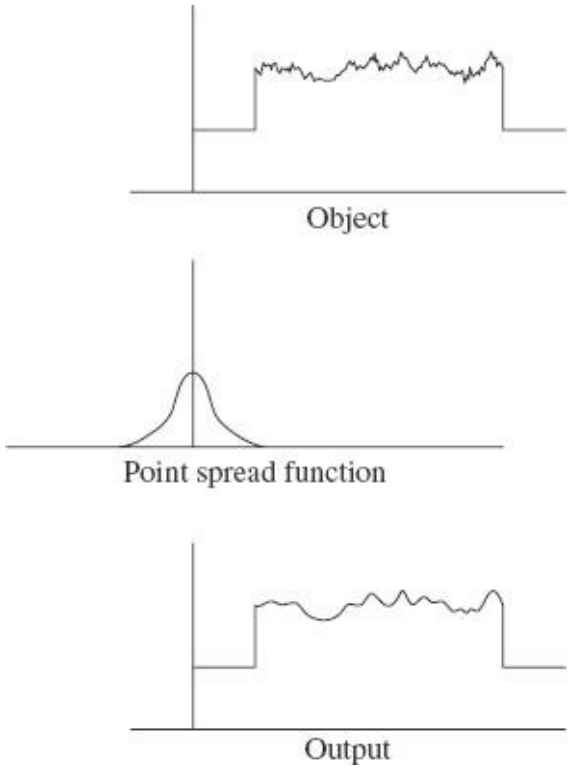


FIGURE 15-1 The effect of convolution of an ideal image with the point spread function.

While there are techniques known as *optimal filters* that can negate these detrimental effects to some extent, their implementation is complex and requires an in-depth knowledge of signal

processing theory. On the other hand, simpler methods are also available which can at least partially compensate for these effects. Some of these methods can reduce image noise, and others can reduce the blurring caused by the point-spread function through edge enhancement.

15-3 Spatial Frequency of a Digital Image

Section 2-12 gave a general description of digital images and described the processes by which they are produced. It introduced the concepts of pixels as discrete samples of the image at a specific geometric resolution, as well as the quantization of brightness values. These are important concepts in understanding digital image processing. Another notion that is key to the comprehension of digital image processing is that of *spatial frequency* (see Sec. 3-14). Objects (e.g., terrain features) have a virtually unlimited range of spatial frequencies. Imagine, for example, an ideal image of a cornfield, where the rows are nominally spaced 1 m apart. On a small-scale image (e.g., high-altitude air photo), the variations from light to dark to light, etc. (corresponding to corn, soil, corn, etc.) might be barely discernible. The spatial frequency at object scale associated with this variation would be 1 cycle per meter. A medium-scale image (e.g., low-altitude air photo) of the same cornfield might reveal variations from light to dark between individual leaves of the plants. If the leaves were nominally spaced 10 cm (0.1 m) apart, the corresponding spatial frequency would be 1 cycle per 0.1 m, or 10 cycles per meter. A large-scale image (e.g., close-up photo taken with a handheld camera) might reveal variations associated with individual kernels on an ear of corn. If the kernels were nominally spaced 5 mm (0.005 m) apart, the corresponding spatial frequency would be 1 cycles per 0.005 m, or 200 cycles per meter. If an individual kernel were viewed under a microscope, even higher spatial frequencies would be discernible. The point of this example is that a multitude of different frequencies exists in objects in nature.

Although there are many different frequencies in a natural scene, most of the higher frequencies will be lost when a digital image is acquired. This occurs whether the image is exposed directly by a digital sensor or is obtained by scanning a film-based photograph. In the process of obtaining digital images, brightness variations that occur within an individual pixel will be averaged together, thus attenuating high frequencies. This is a consequence of the physical size of the pixel being sampled. The highest frequency that can be represented in a digital image is the *Nyquist frequency*, which is one-half the sampling frequency. As an example, in a satellite image with pixels having a 10-m ground resolution, the sampling frequency is 1 sample per 10 m, or 0.1 sample per meter. The highest spatial frequency (Nyquist frequency) that can be represented in this image is therefore 0.05 cycles per meter, or 1 cycle per 20 m. Note that 1 cycle corresponds to two samples. Figure 15-2 illustrates the effect of using different sample frequencies to represent a given signal. The signal frequency is eight cycles per unit. Notice that a sampling frequency of 16 samples per unit is the minimum to fully represent the signal. The Nyquist frequency is the upper limit for the modulation transfer function (see Sec. 3-13). This is illustrated in Fig. 15-3. Since the sampling frequency in the figure is 160 samples per mm, the modulation transfer function is not applicable above the Nyquist frequency, 80 cycles per mm.

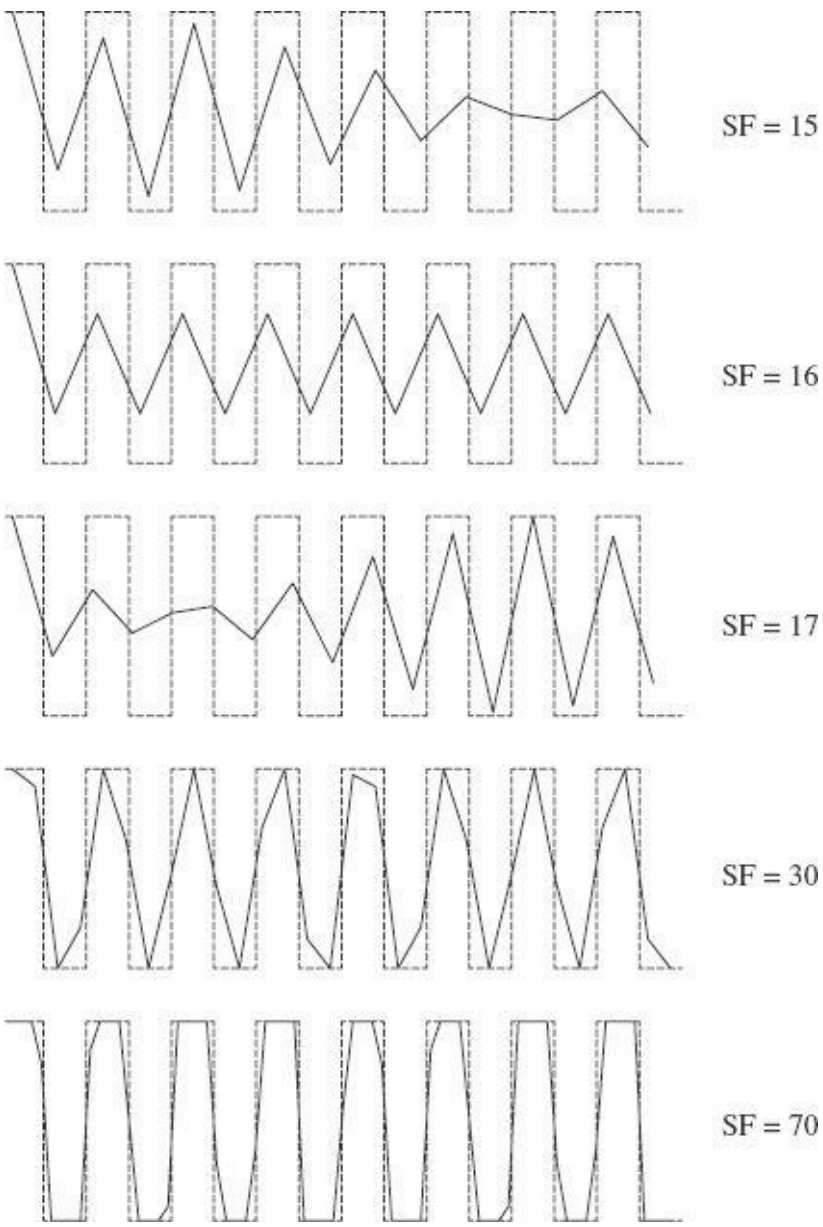


FIGURE 15-2 The effect of different sampling frequencies in the depiction of a signal with a frequency of eight cycles per second.

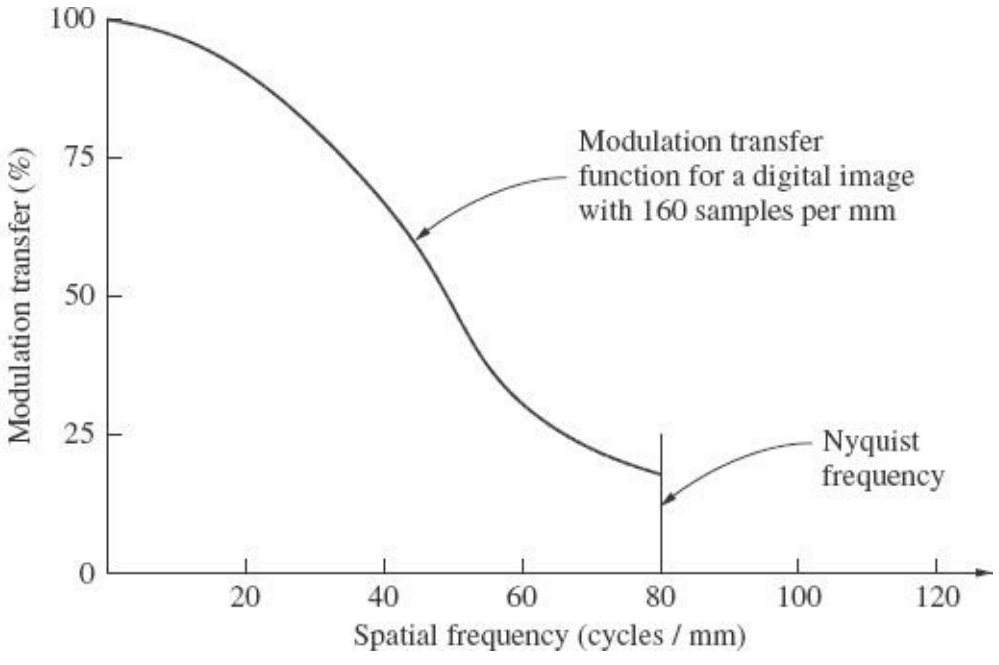


FIGURE 15-3 Relationship between the Nyquist frequency and the modulation transfer function.

The relationship between the variations of digital numbers in a digital image and the spatial frequencies they represent can be precisely quantified in a mathematical sense. Often, certain characteristics of a digital image are more logically described in terms of spatial frequency than by digital numbers in an image. In sections that follow, explanations will be given which should help to clarify this relationship. However, a full understanding of the concept will require substantial study in the area of signal processing theory, which is beyond the scope of this text.

15-4 Contrast Enhancement

Contrast enhancement is a commonly employed type of digital image processing operation. As discussed in Sec. 2-12, when each pixel of a scene is quantized, a digital number that corresponds to the brightness at its location is assigned. In an area where the overall scene brightness is low (e.g., a shadow region), the majority of the pixels will have low values. In such cases, the contrast between features will be diminished. To get an impression of the overall brightness values in an image, it is useful to produce a *histogram* of the digital numbers. The histogram is a plot of digital number on the abscissa versus number of occurrences on the ordinate. As an example, if a digital image has 1907 pixels having a digital number (DN) of 29, the histogram will show an ordinate of 1907 at the abscissa of 29. Figure 15-4*a* shows a small portion of a digital image which is in the shadow of a building, and Fig. 15-4*b* shows its corresponding histogram of digital numbers. The image contains a sidewalk and a bench which are barely noticeable because of the low contrast caused by the shadow. Notice from the histogram that the majority of digital numbers falls in the range of 20 to 65, even though the available range is 0 to 255. This indicates that when the image is displayed, the contrast among features in this area will be low, as is certainly evident in Fig. 15-4*a*.

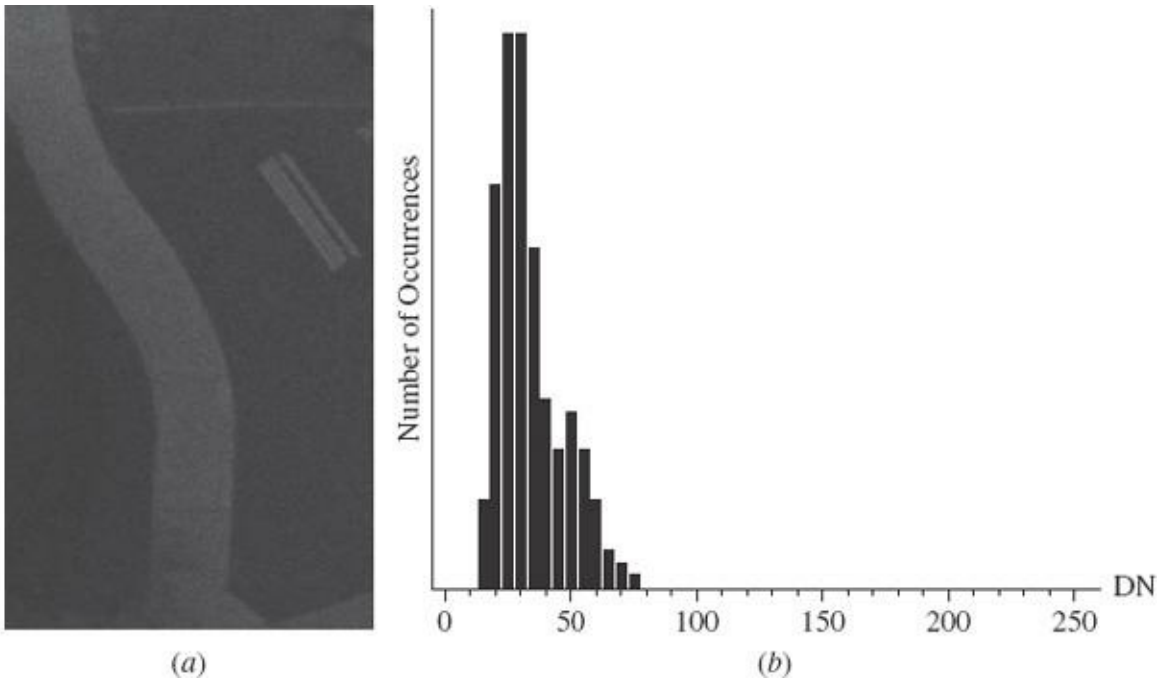


FIGURE 15-4 (*a*) Digital image of an area in a shadow. (*b*) Histogram of digital numbers from the image.

A simple method of increasing the contrast is to apply a *linear stretch*. With this operation, the digital numbers within a certain range (e.g., from 20 to 65) are linearly expanded to the full available range of values (e.g., from 0 to 255). The effect of a linear stretch on the image of Fig. 15-4*a* is illustrated in Fig. 15-5*a*. Its associated histogram is shown in Fig. 15-5*b*. Notice in this histogram that the digital numbers have been expanded (stretched) so as to encompass the full available range (0 to 255). As can be seen in Fig. 15-4*a*, the effect of this linear stretch is enhanced contrast, making the sidewalk and bench more apparent.

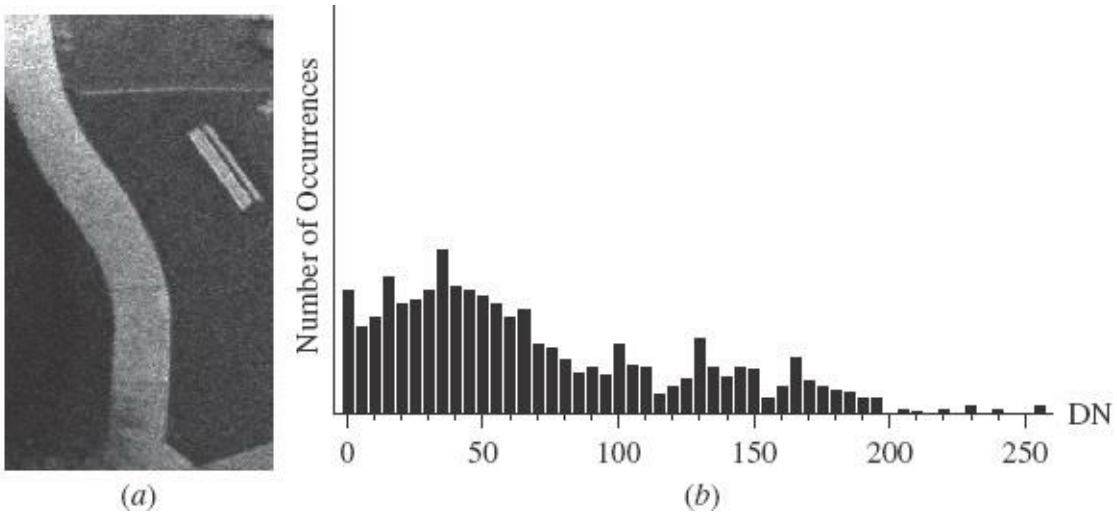


FIGURE 15-5 (*a*) Digital image from Fig. 15-4*a* after linear stretch contrast enhancement. (*b*) Histogram of digital numbers from the image after linear stretch.

A more complex method of increasing the contrast is to apply *histogram equalization*. The concept behind histogram equalization is to expand the digital numbers in a nonlinear fashion across the available range so that the values are more evenly distributed. The effect of histogram equalization on the image of Fig. 15-4*a* is illustrated in Fig. 15-6*a*. Its associated histogram is shown in Fig. 15-6*b*. Notice from this histogram that the numbers are more evenly distributed, particularly at the high end of the digital number range. Note also that the histogram-equalized image of Fig. 15-6*a* has even greater contrast than that of the linear stretched image, making the sidewalk and bench even more pronounced.

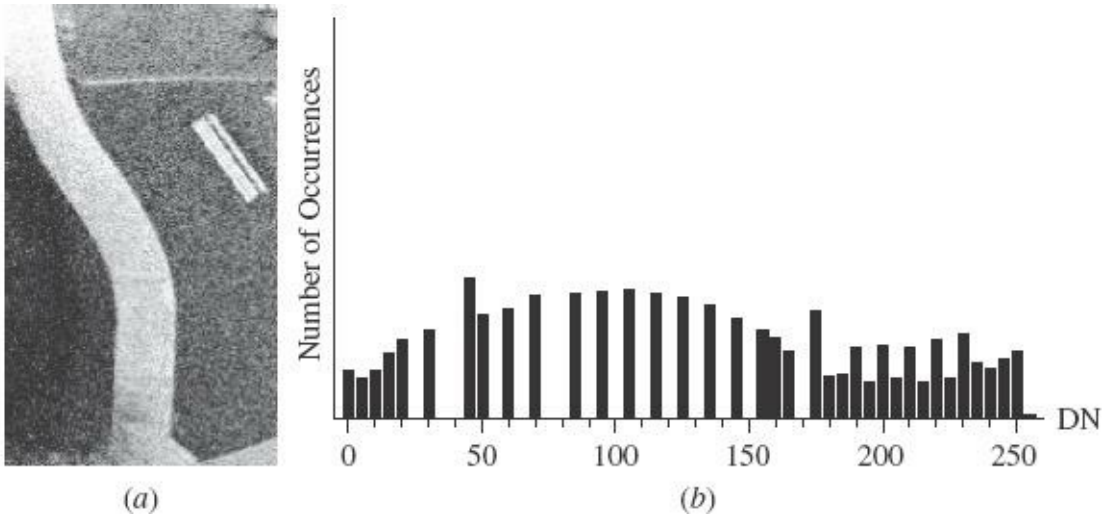


FIGURE 15-6 (*a*) Digital image from Fig. 15-4*a* after histogram equalization contrast enhancement.

(*b*) Histogram of digital numbers from the image after histogram equalization.

15-5 Spectral Transformations

As discussed in Sec. 15-3, digital images capture information at various spatial frequencies. Digital images in their native form exist in the *spatial domain*. The term *spatial domain* refers to the concept that the coordinates of pixels in the image relate directly to positions in the two-dimensional space of the image. Through certain mathematical operations, images can be converted from the spatial domain to the *frequency domain*, and vice versa. A digital image which has been converted to the frequency domain has pixels whose positions (row and column) relate to frequency instead of spatial locations. Images in the frequency domain no longer have visibly recognizable features as spatial domain images do. They are an abstraction of the original spatial image; however, they do contain the entire information content of the original. In the frequency domain, the spatial frequencies inherent in the native image are represented by measures of amplitude and phase for the various frequencies present in the image.

The most commonly used method for converting a digital image to the frequency domain (or vice versa) is the *discrete Fourier transform* (DFT). This process is roughly analogous to conversion between geodetic and geocentric coordinates (see Sec. F-3). Converting from geodetic to geocentric coordinates changes the representation of a point's position from latitude, longitude, and height; to *X*, *Y*, and *Z*. In an analogous fashion, the DFT converts the brightness values from a (spatial domain) digital image to a set of coefficients for sine and cosine functions at frequencies ranging from 0 to the Nyquist frequency. In an additional step, amplitude and phase information can be obtained from these sine and cosine coefficients. A more detailed explanation of the DFT requires the use of complex numbers, which is beyond the scope of this text.

The discrete Fourier transform has become a useful tool for analyzing digitally sampled periodic functions in one, two, or even more dimensions. (Note that, as described in Sec. 15-2, a digital image is a two-dimensional function.) A particularly efficient implementation of the discrete Fourier transform is the aptly named *fast Fourier transform* (FFT). Since its discovery and implementation in the 1960s, it has enabled discrete Fourier transforms to be computed at a greatly enhanced speed. This greatly enhanced computational speed has led to wider use of the Fourier transform as an image processing tool.

Figure 15-7*a* shows a digital satellite image of a portion of Gainesville, Florida, with its Fourier transform (amplitudes only) shown in Fig. 15-7*b*. In Fig. 15-7*b*, greater amplitudes correspond to brighter tones, and lesser amplitudes correspond to darker tones. This Fourier transform exhibits a star-shaped pattern which is typical of two-dimensional transforms of urban scenes. The lowest spatial frequencies are represented at the center of the star with frequency increasing with increasing distance from the center. It can be seen from Fig. 15-7*b* that lower frequencies have greater magnitudes than higher frequencies. The figure also shows discernible spikes radiating from the center. Those radiating in vertical and horizontal directions are due to discontinuities associated with the edges of the image, whereas the spikes oriented approximately 10° counterclockwise are due to the edges associated with the primary streets and buildings in the image.

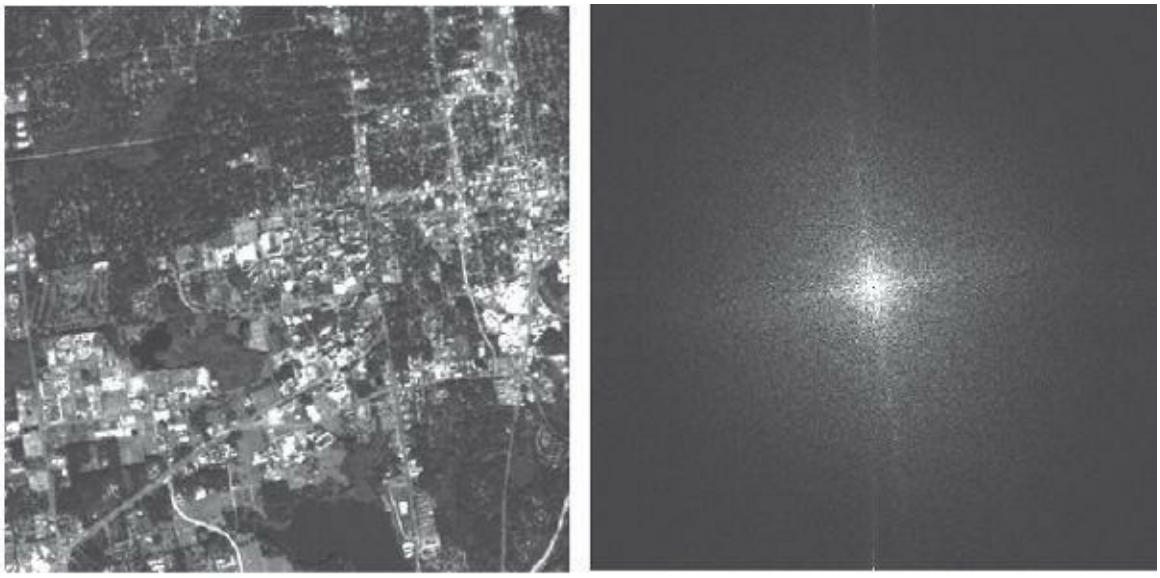


FIGURE 15-7 (*a*) Digital satellite image of Gainesville, Florida. (*b*) Fourier transform. (*Courtesy University of Florida.*)

Certain operations are easier to accomplish in the frequency domain than in the spatial domain. For example, assume a digital sensor had an electrical problem which caused an interference pattern in the image as illustrated in Fig. 15-8*a*. Conversion of this contaminated image from the spatial to the frequency domain reveals high amplitudes for the particular frequencies which contributed to the interference, as shown in Fig. 15-8*b*. The coefficients corresponding to the sine and cosine terms for those frequencies can be changed from the high values to zero values as illustrated in Fig. 15-8*c*, thus eliminating the interference in the frequency domain image. To view the image, an inverse Fourier transform would then be applied to the modified frequency image to convert back to the spatial domain, as illustrated in Fig. 15-8*d*. Since the systematic interference was eliminated in the frequency domain, it is no longer present in the spatial domain image. The interference removal was not perfect, however, as can be seen in the subtle interference patterns which remain in Fig. 15-8*d*.

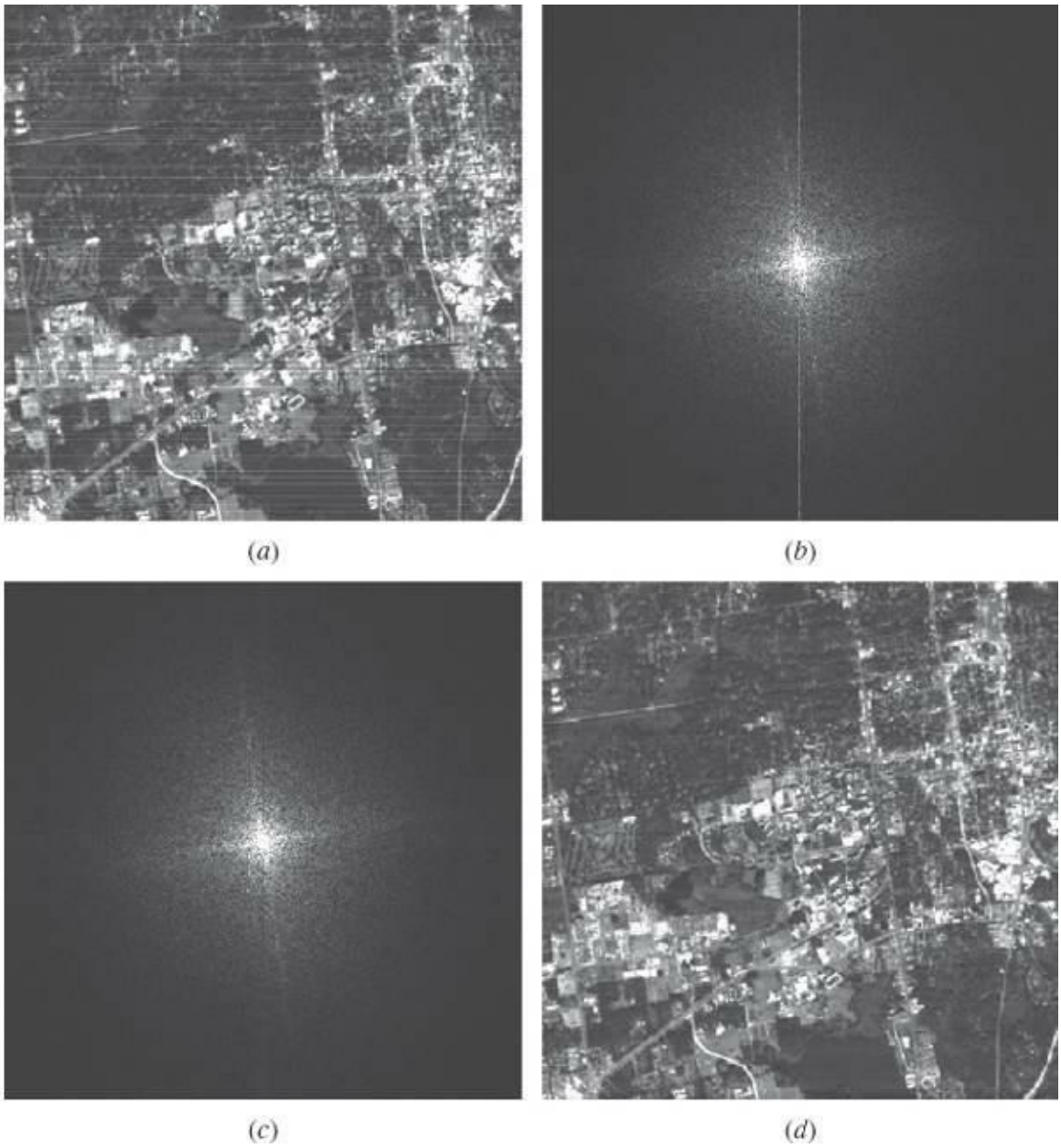


FIGURE 15-8 (*a*) Gainesville image with interference noise. (*b*) Fourier transform. (*c*) Fourier transform with high-frequency interference eliminated. (*d*) Cleaned image after inverse Fourier transform of (*c*). (*Courtesy University of Florida.*)

As mentioned above, certain mathematical operations are easier to compute in the frequency domain than the spatial domain. Another such operation is *convolution*. This process involves two functions: the signal (e.g., digital image) and the response function. An example of a response function is the mathematical representation of the effect that a real imaging system imparts to an ideal image to produce the actual image (i.e., the point-spread function). The two functions are convolved, which will apply the "smearing" effect of the response function to the signal. Depending on the form of the response function, the convolution will have the effect of filtering out certain high frequencies from the image. In the absence of noise, the inverse operation *deconvolution* can be applied to an image that has been affected by a known response function, to reconstruct the high frequencies which were lost. Deconvolution may be done to remove the blurring effect of the point-spread function (see Sec. 15-2) in order to recover the ideal image from an imperfect real image. Deconvolution should be done with caution, however, since it can be highly sensitive to image noise.

Another operation which is easy to compute in the frequency domain is *correlation*. Correlation also involves two functions, which can both be assumed to be signals (e.g., digital images).

Correlation provides a measure of the similarity of two images by comparing them while superimposed, as well as shifting an integral number of pixels left, right, up, or down, all with one simultaneous operation. Correlation can be useful in matching patterns from one image that correspond to those of another image. Pattern matching provides the backbone of softcopy mapping systems.

Fourier transforms, while probably the best known, are not the only image transformations available. Another transformation is the *wavelet transform.* Wavelets are types of functions characterized by an oscillation over a fixed interval. They can be used to represent an image in a similar manner as the Fourier transform. In practice, a “mother wavelet” serves as a template for several “child” wavelets, which are scaled and translated to match the frequencies in the rows and columns of an image. Since all that is needed to represent an image are the scale and translation parameters, wavelets can reduce the amount of storage space for an image by eliminating the coefficients of rarely used frequencies. Thus, wavelets are useful in image compression.

Another useful transformation is the Hough transform. The Hough transform is primarily used for joining discontinuous edges in an image. It can be highly useful for reconstructing linear and circular features from raster images, where the edges have breaks in them. Details of these transforms are beyond the scope of this text; however, references cited at the end of this chapter can be consulted for further information.

Figure 15-9 shows a first-level wavelet decomposition. In the upper left corner is an image at half its original resolution (low frequency components). The three other images contain the high frequency components removed when decreasing the resolution. The upper-right hand image contains the high frequencies components in the horizontal direction, the lower-left hand image contains the components associated with high frequencies in the vertical direction, and the lower-right hand image contains the components for diagonal high frequencies. These four sub-images can be used to perfectly recreate the image in its original resolution. The full wavelet transformation entails iteratively reducing the resolution of previous-level images and storing the higher frequency components until the lower resolution image is one pixel. Similar to the Fourier transform, wavelets also allow the smoothing of imagery by withholding unwanted frequencies when recreating an image. Wavelets can also be used in many other image processing algorithms such as edge detection.

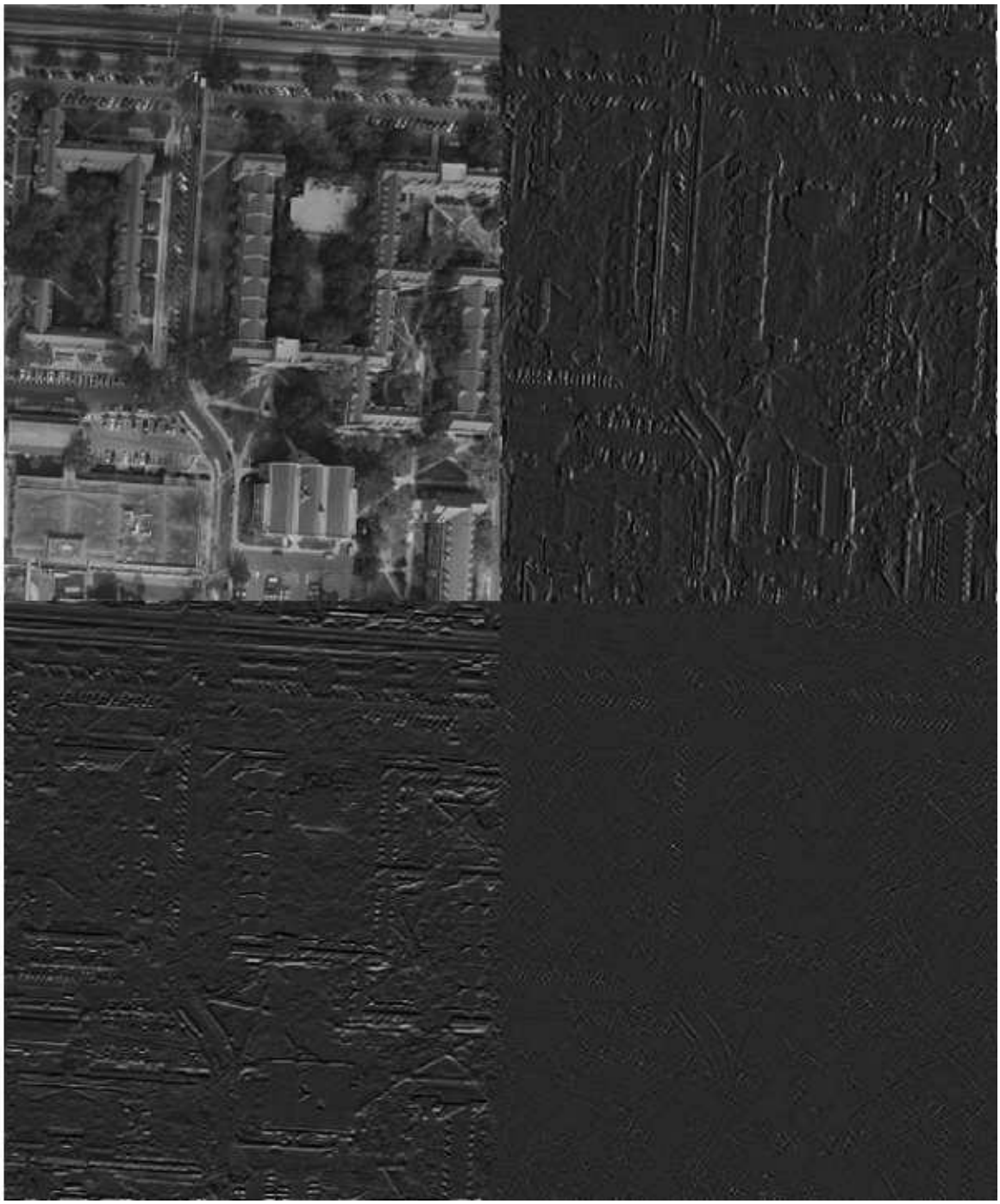


FIGURE 15-9 The first level of wavelet decomposition for an aerial image.

15-6 Moving Window Operations

As mentioned in Sec. 15-5, the Fourier transform is highly effective for performing convolution of a digital image with a response function. A simpler method known as the *moving window* can also be used for performing convolution. The moving window is particularly useful when the response function is highly localized. For example, a two-dimensional response function may consist of a 3×3 submatrix of nonzero terms contained within a large array which has zero terms everywhere else. When convolving using the Fourier transform, the image and response function must have the same dimensions, which can lead to inefficiency when the vast majority of terms in the response function are equal to zero. In cases such as this, the moving window can be used efficiently.

The moving window operation has two primary inputs: the original digital image and a localized response function known as a *kernel*. The kernel consists of a set of numbers in a small array, which usually has an odd number of rows and columns. The kernel is conceptually overlaid with a same-size

portion of the input image (image window), and a specific mathematical operation is carried out. The value resulting from this operation is placed in the center of the corresponding location in the output image, the kernel is shifted by one column (or row), and another value is computed. This procedure is repeated until the entire output image has been generated. Figure 15-10 illustrates one step of the procedure. In this figure, the 3 × 3 kernel is centered on pixel 3, 4 (row = 3, column = 4) in the input image. After the convolution is performed, the result is placed in pixel 3, 4 of the output (convolved) image. The kernel is then shifted so that it is centered on pixel 3, 5, and the convolution is repeated. When convolving pixels at the edge of the original image, the kernel will extend outside the image boundary. Under the assumption that the image is a periodic function, the kernel values that extend beyond the image will "wraparound" so as to overlay with image pixels at the opposite side of the image. The wrap-around effect can be visualized by imagining exact copies of the image to exist above, to the right, below, to the left, and diagonally away from each of the four corners of the original image. This is the same result that would have occurred if the Fourier transform had been used to perform the convolution.

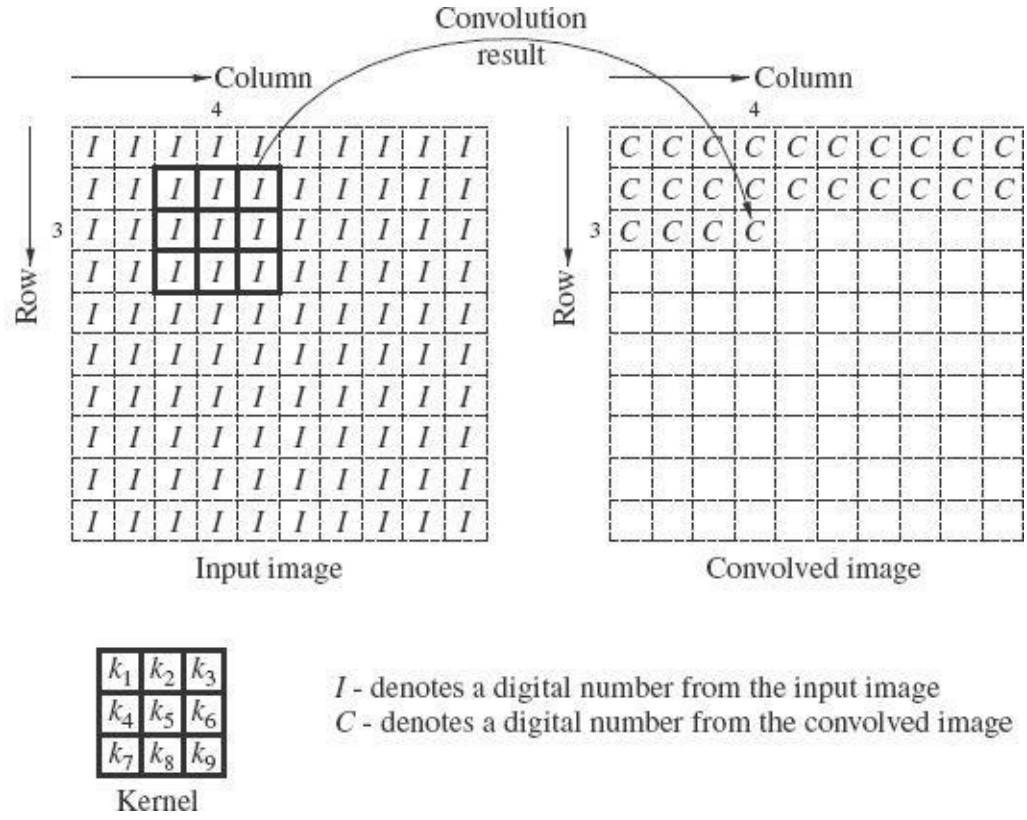


FIGURE 15-10 Moving-window image convolution.

The mathematical operation of convolution is accomplished through a series of multiplications and additions. For example, the convolution result depicted in Fig. 15-10 is computed by

$$C_{34} = k_1 \times I_{23} + k_2 \times I_{24} + k_3 \times I_{25} + k_4 \times I_{33} + k_5 \times I_{34} + k_6 \times I_{35} + k_7 \times I_{43} + k_8 \times I_{44} + k_9 \times I_{45} \quad (15\text{-}2)$$

In Eq. (15-2), the k values are the individual kernel elements, the I values are digital numbers from the input image with the appropriate row and column indicated by their subscripts, and C_{34} is the convolution result which is placed at row = 3, column = 4 in the output image.

Perhaps the simplest form of convolution kernel is one that computes the average of the nearby pixels. This type of convolution is known as a *low-pass filter*, so called because the averaging operation attenuates high frequencies, allowing low frequencies to be passed on to the convolved image. The kernel values used in a simple 3 × 3 low-pass filter are shown in Eq. (15-3).

$$K = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix} \tag{15-3}$$

In this equation, K is the kernel matrix, which in this case has all elements equal to $\frac{1}{9}$. Convolution of this kernel matrix with a corresponding 3 × 3 image submatrix is equivalent to computing a weighted average of the digital numbers within the submatrix. Notice that the sum of the elements in this kernel matrix equals 1, which is appropriate when computing a weighted average. In fact, elements of a kernel can be considered to be the weights used in computing a weighted average. If the sum of the elements of the kernel matrix does not equal 1, the convolution result can be divided by the kernel's sum before placement into the output image.

Example 15-1

Given the following convolution kernel K and the matrix of digital numbers D from a 5 × 5 image, compute the convolved image C, using the moving window approach.

$$K = \begin{bmatrix} 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.1 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$

$$D = \begin{bmatrix} 84 & 79 & 43 & 74 & 76 \\ 81 & 77 & 45 & 75 & 78 \\ 50 & 48 & 39 & 42 & 44 \\ 80 & 77 & 46 & 73 & 72 \\ 82 & 83 & 47 & 71 & 70 \end{bmatrix}$$

Solution

1. Create the periodic wrap-around effect by adding an extra row and column from the imaginary copies of the image surrounding the original image.

$$D' = \begin{bmatrix} 70 & 82 & 83 & 47 & 71 & 70 & 82 \\ 76 & 84 & 79 & 43 & 74 & 76 & 84 \\ 78 & 81 & 77 & 45 & 75 & 78 & 81 \\ 44 & 50 & 48 & 39 & 42 & 44 & 50 \\ 72 & 80 & 77 & 46 & 73 & 72 & 80 \\ 70 & 82 & 83 & 47 & 71 & 70 & 82 \\ 76 & 84 & 79 & 43 & 74 & 76 & 84 \end{bmatrix}$$

2. Convolve at each 3 × 3 position of the moving window. Calculation of the first two elements will be shown as examples.

$$c_{11} = 0.1 \times 70 + 0.1 \times 82 + 0.1 \times 83 + 0.1 \times 76 + 0.2 \times 84 + 0.1 \times 79 + 0.1 \times 78 + 0.1 \times 81 + 0.1 \times 77$$

$$c_{11} = 79.4 = 79 \text{ (rounded to nearest integer)}$$

$$c_{12} = 0.1 \times 82 + 0.1 \times 83 + 0.1 \times 47 + 0.1 \times 84 + 0.2 \times 79 + 0.1 \times 43 + 0.1 \times 81 + 0.1 \times 77 + 0.1 \times 45$$

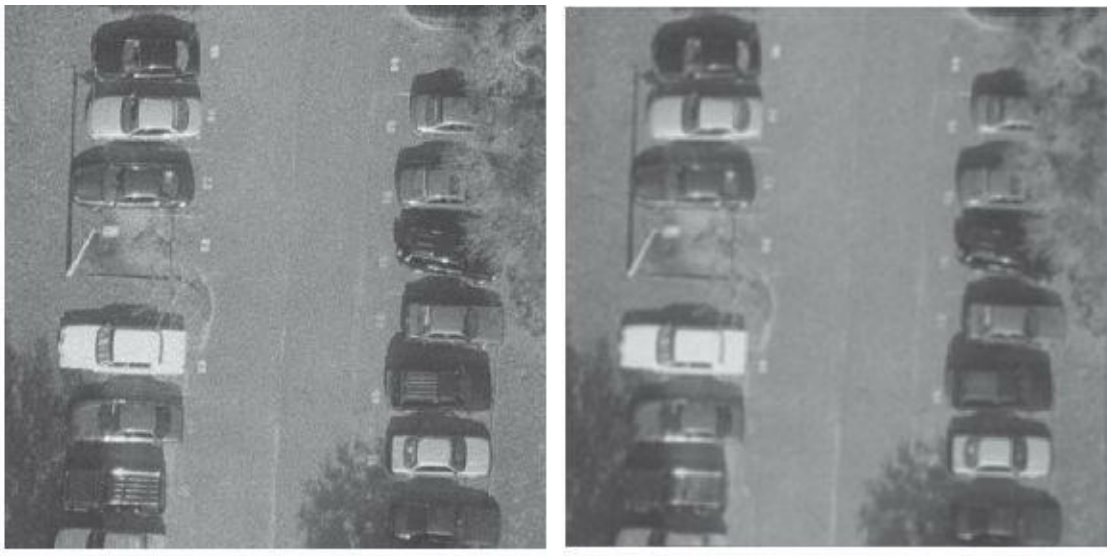
$$= 70.0 = 70 \text{ (rounded to nearest integer)}$$

the remaining values are

$$C = \begin{bmatrix} 79 & 70 & 64 & 65 & 77 \\ 70 & 62 & 57 & 59 & 68 \\ 66 & 59 & 56 & 56 & 64 \\ 69 & 63 & 57 & 58 & 66 \\ 79 & 70 & 64 & 64 & 75 \end{bmatrix}$$

▲

The result of a low-pass filter convolution is illustrated in Figs. 15-11*a* and *b*. Figure 15-11*a* shows the original image of cars in a parking lot. A low-pass filter was applied using a 5 × 5 convolution kernel which computed a simple average, resulting in the image of Fig. 15-11*b*. Notice that the image of Fig. 15-11*b* now appears somewhat blurred since the highest-frequency detail has been filtered out from the original image. By applying kernels of larger size (say, 7 × 7 or 9 × 9) a wider range of high-frequency information will be filtered from the image. The ultimate limit in kernel size would be the same as that of the original image. Such a convolution would result in all pixels of the output image having the same value—the overall average of digital numbers from the input image. In that case, detail at all frequencies will have been filtered out, leaving a uniformly gray image.



(*a*) (*b*)

FIGURE 15-11 (*a*) Parking lot image. (*b*) Parking lot image after applying a low-pass filter.

The intentional blurring of an image as the result of applying a low-pass filter may seem like a counterproductive operation. After all, high-frequency detail assists in identifying features and measuring the positions of their edges. The method could serve as a simple high-frequency noise

filter, but there are better methods available for this purpose. However, another use for low-pass filtering is to employ the process as a precursor to a *high-pass* operation. By performing a pixel-by-pixel subtraction of the low-passed image from the original image, the result will give the high-frequency detail which was filtered out in the low-pass operation. Since subtraction can yield negative numbers which are inconvenient to deal with in a digital image, the absolute value can be taken after each subtraction. Figure 15-12 shows the high-pass filtered result from subtracting the digital numbers of Fig. 15-11*b* from those of Fig. 15-11*a* and taking the absolute value. As is apparent from the figure, this simple high-pass filter can be used to detect edges in the original image.

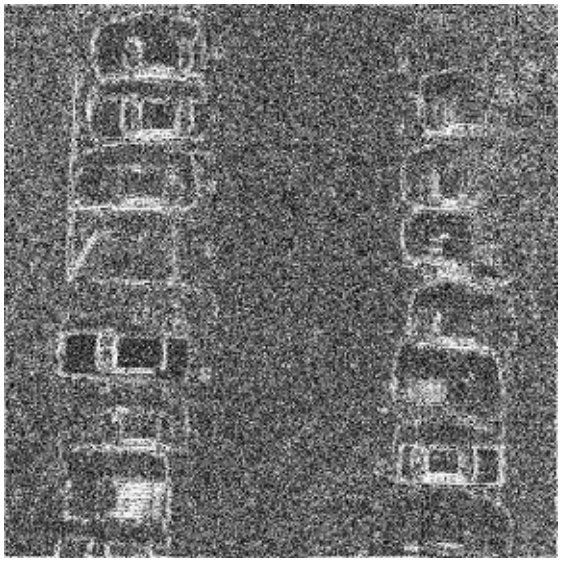


FIGURE 15-12 Parking lot image after application of a simple high-pass filter.

Moving window operations can be used to perform other useful operations, such as noise filtering. One such noise filter method is known as *median filtering*. In this approach, a convolution is not performed in the normal fashion. Rather, a 3 × 3 moving window is passed through the input image, and the 9 pixels in the immediate neighborhood are extracted at each step. The nine digital numbers are then sorted, and the median (middle) value is placed in the corresponding location in the output image. By using the middle value rather than the average, the median filter will not be sensitive to any extremely high or low value which may be the result of image noise. The result of median filtering is shown in Figs. 15-13*a* and *b*. Figure 15-13*a* shows the same image of a parking lot shown in Fig. 15-11*a*, except that for purposes of this example, random "salt and pepper" noise has been added. After the median filter operation is applied, the image of Fig. 15-13*b* results. Note that the noise has indeed been eliminated, at the expense of a subtle loss of high-frequency information. Median filters are useful for removing many forms of random noise, but are not as effective at removing systematic noise.

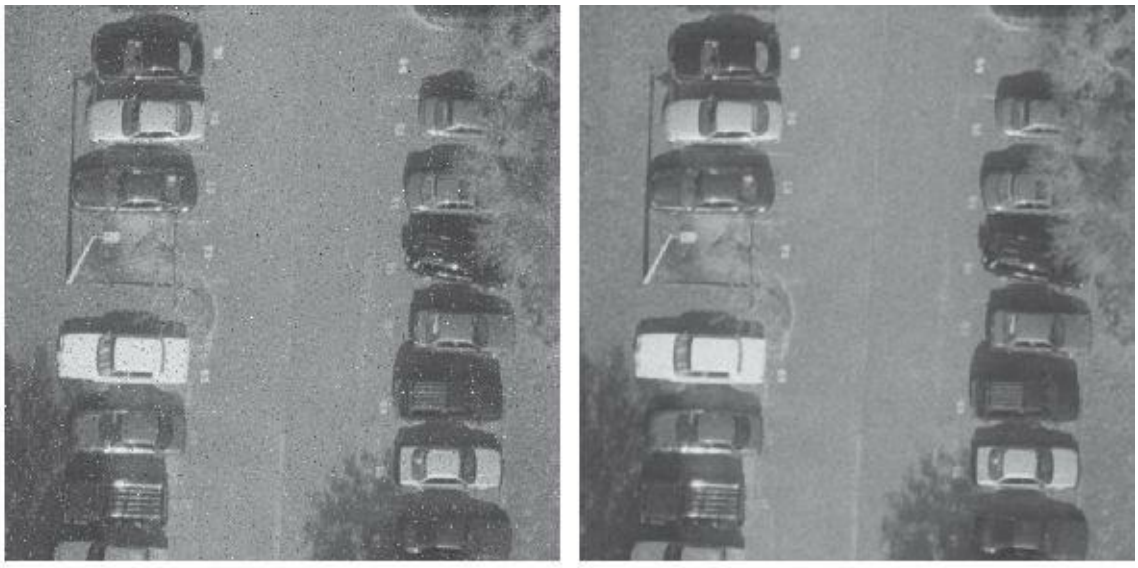


FIGURE 15-13 (*a*) Parking lot image with noise. (*b*) Parking lot image after application of median filter.

Another class of operations which uses the moving window approach is edge detection. Earlier in this section, a simple form of edge detection based on high-pass filtering was presented. Other methods exist for edge detection, and two specific ones are presented here: the laplacian and Sobel operators.

The laplacian operator is a convolution using a kernel which has values corresponding to the shape of the function shown in Fig. 15-14, except that it extends in two dimensions. In Fig. 15-14, note that at the center, the laplacian starts with a positive value and drops off to negative values at a certain distance from the center. The effect of this type of kernel is to amplify the differences between neighboring digital numbers which occur at an edge. If the kernel values (weights) are chosen so that their sum is equal to zero, the laplacian will result in a high-pass operation, with the lower frequency information filtered out. The result will be an image which contains edge information at the corresponding high frequency. By using a larger laplacian kernel which is flatter and wider, edge information at a particular frequency will be passed to the output image, resulting in a *bandpass* image. A bandpass image is an image which contains frequencies within a specific range only. Frequencies outside the range (band) are not present in the bandpass image. Note that since the laplacian kernel contains negative values, the convolution result can be negative, so the absolute value should be taken if the result is to be displayed.

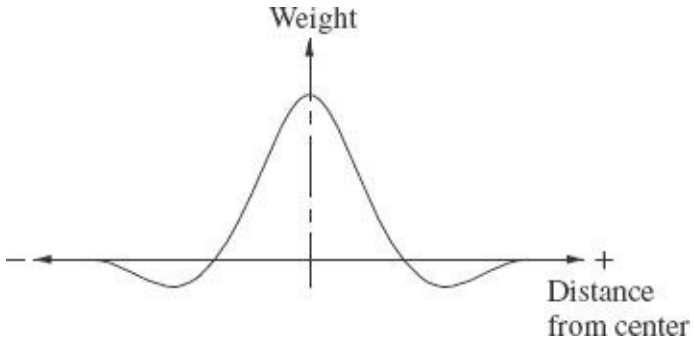
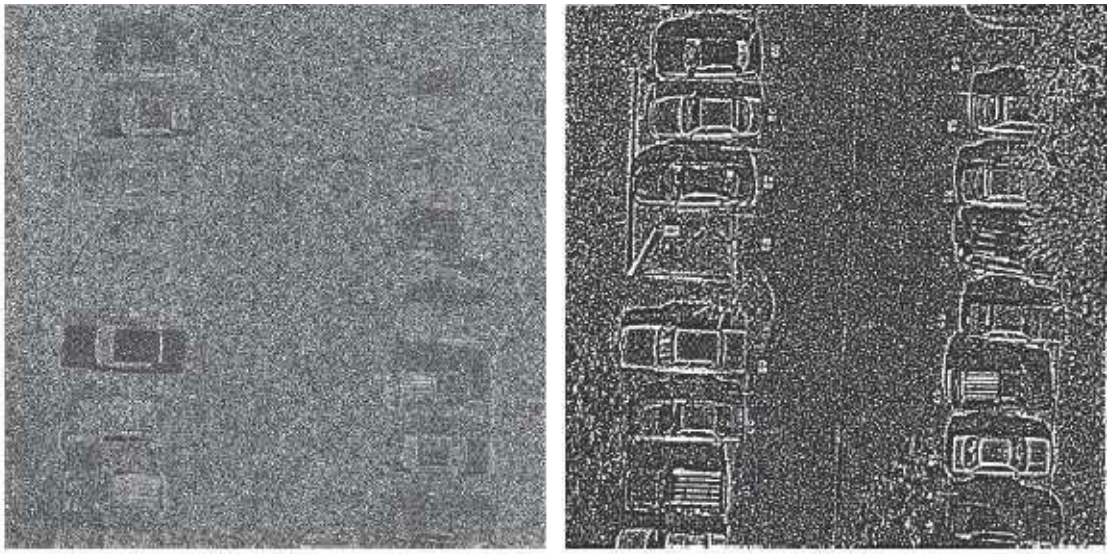


FIGURE 15-14 Laplacian weighting function.

The results of two laplacian convolutions on the parking lot image of Fig. 15-11*a* are shown in Figs. 15-15*a* and *b*. The convolution kernel used for Fig. 15-15*a* has the form shown in Eq. (15-4), and the kernel used for Fig. 15-15*b* has the form shown in Eq. (15-5). Note that the sum of the elements for each of these kernels is equal to zero.

$$K = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{15-4}$$



(*a*) (*b*)

FIGURE 15-15 (*a*) Parking lot image after convolution with 3 × 3 laplacian kernel. (*b*) Parking lot image after convolution with 5 × 5 laplacian kernel.

$$K = \begin{bmatrix} -1 & -2 & -3 & -2 & -1 \\ -2 & 2 & 4 & 2 & -2 \\ -3 & 4 & 8 & 4 & -3 \\ -2 & 2 & 4 & 2 & -2 \\ -1 & -2 & -3 & -2 & -1 \end{bmatrix} \tag{15-5}$$

Notice from Fig. 15-15*a* that the laplacian operation has enhanced the highest frequencies while filtering out all other frequencies. The resulting edges correspond to the highest frequencies and barely contain recognizable information. Since a wider laplacian kernel was used to produce Fig. 15-15*b*, the edges that were detected correspond to frequencies that are slightly lower than those of Fig. 15-15*a*. Notice that the outlines of the cars are readily discernible, as well as the lamppost and its shadow.

The laplacian edge detector is not sensitive to the orientation of the edge that it detects. A different edge detection operator, the Sobel, is capable of not only detecting edges, but determining their orientation as well. This is accomplished by using a pair of kernels (an *x* kernel and a *y* kernel) in the convolution. The forms of the *x* and *y* kernels are given in Eqs. (15-6) and (15-7), respectively.

$$K_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \tag{15-6}$$

$$K_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{15-7}$$

Convolution by the two kernels at a specific window location gives the x component of the edge, S_x, and the y component of the edge, S_y. From these two convolution results, the magnitude and direction of the edge can be computed by Eqs. (15-8) and (15-9), respectively.

$$\text{Magnitude} = \sqrt{S_x^2 + S_y^2} \tag{15-8}$$

$$\text{Direction} = \tan^{-1}\left(\frac{S_y}{S_x}\right) \tag{15-9}$$

Application of the Sobel edge detection operation to the parking lot image results in the magnitude image shown in Fig. 15-16. Notice that the edges defining the cars and their shadows are well defined, as are the edges defining the lamppost and its shadow.

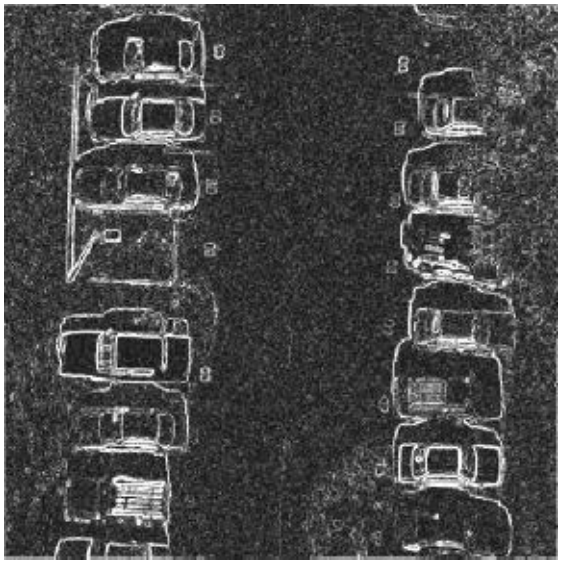


FIGURE 15-16 Parking lot image after application of Sobel edge detection.

From the examples given on this and the previous page, it can be seen that the size and composition of the kernels applied in digital image processing have a major bearing upon the results achieved, and they also create many interesting possibilities.

15-7 Multiscale Representation

In the preceding section, bandpass filters were discussed. Images that result from a bandpass filter contain edge information that corresponds only to the range of frequencies in the bandpass image. By using filters which pass lower ranges of frequencies, low-frequency (small-scale) edges can be detected. The effect of detecting edges at different scales (frequencies) can be seen in Figs. 15-17*a*, *b*, and *c*. Figure 15-17*a* is the result of performing a simple average convolution with a 3 × 3 kernel on the parking lot image, and subtracting the convolved image from the original. Note that this image contains a great deal of large-scale (high-frequency) edge information.

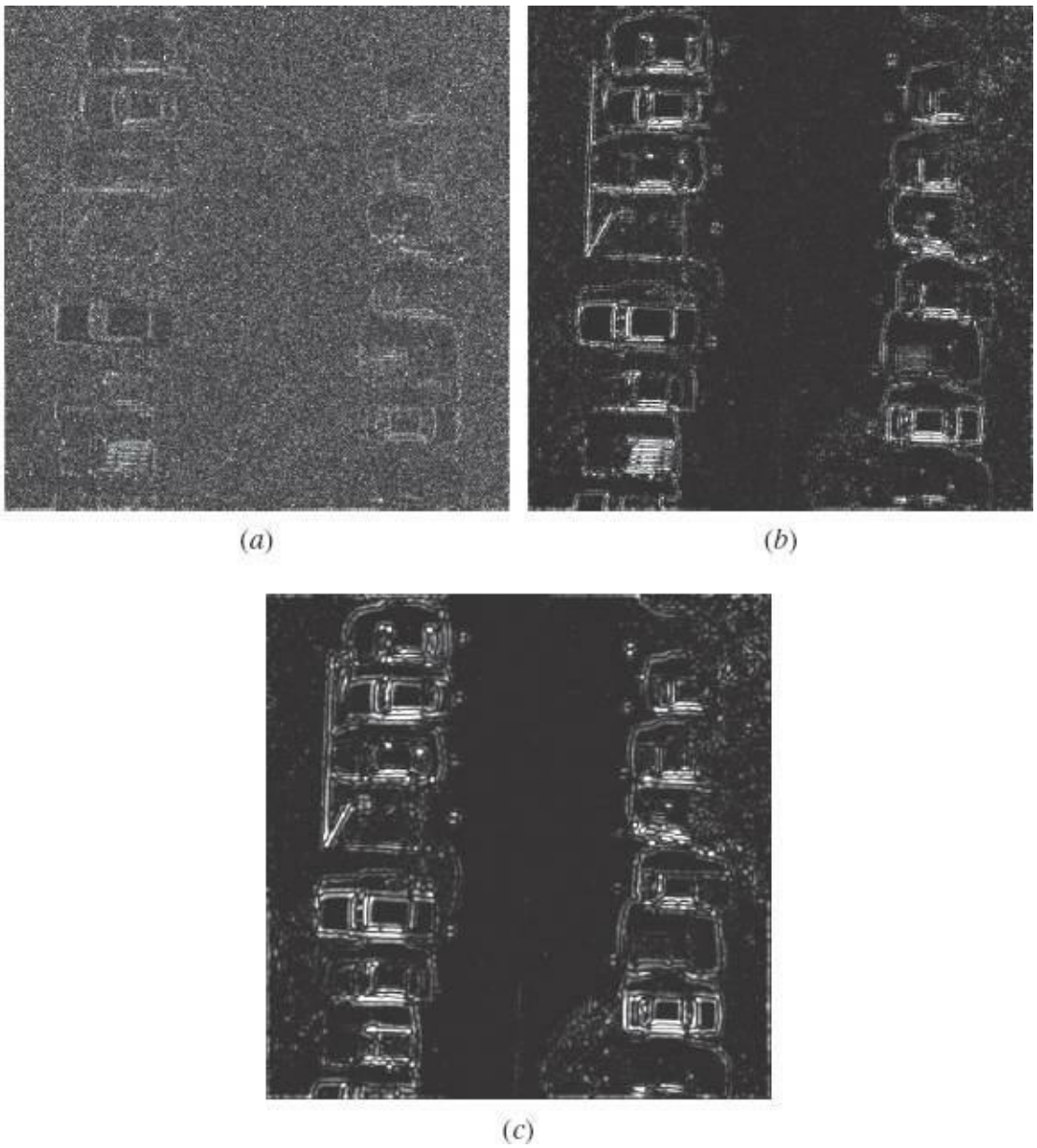


FIGURE 15-17 (*a*) Large-scale edge information from parking lot image. (*b*) Edge information at a smaller scale. (*c*) Edge information at a still smaller scale.

Figure 15-17*b* is the result of two convolutions, followed by a difference. The first convolution is the same as that used to produce Fig. 14-17*a*, i.e., a simple average convolution with a 3 × 3 kernel. Assume that the original parking lot image is image I and the result of the first convolution gives image I_3. The second convolution is performed on image I_3 using a simple average convolution with a 5 × 5 kernel, resulting in image $I_{3,5}$. Finally, image $I_{3,5}$ is subtracted from image I_3, resulting in Fig. 15-

17*b*. Notice that this figure contains edge information at a smaller scale (lower frequency) than that of Fig. 15-17*a*. Notice also that the edges corresponding to the roof rack on the vehicle at the lower left are still discernible.

The last image, Fig. 14-17*c*, is the result of three convolutions followed by a difference. Starting with image $I_{3,5}$ which has been convolved twice, an additional convolution is performed using a 7×7 kernel (simple average), resulting in image $I_{3,5,7}$. Finally, $I_{3,5,7}$ is subtracted from image $I_{3,5}$, resulting in Fig. 15-17*c*. In this figure, edges of a still smaller scale (lower frequency) have been detected. Notice that in this figure, the edges associated with the roof rack have essentially disappeared.

This method of obtaining multiscale representations of edges is rather simplistic, but serves as an introduction to the concept. Other approaches such as those that employ the wavelet transform can produce multiscale representations with better resolution. The wavelet transform is often used to extract multiscale features (edges) for use in feature recognition and pattern matching.

Another concept relating to image scale is that of the *image pyramid*. An image pyramid is formed by successively convolving an image with a *gaussian* kernel, with each convolution producing a halfresolution copy of the previous image. A gaussian kernel is one which has weights that correspond to the shape of a normal (gaussian) distribution or bell-shaped curve. The series of images thus produced can be visualized as a stack of image layers forming a pyramid, as shown in Fig. 15-18. This figure shows an image pyramid formed from the parking lot image, with the original image at the bottom of the pyramid and successive half-resolution copies going up the pyramid.

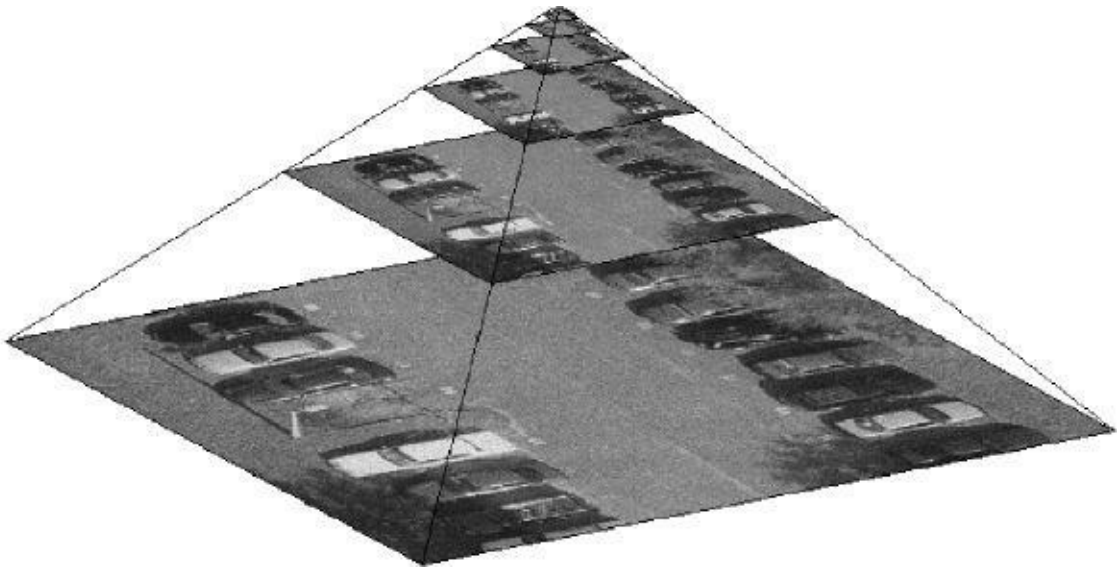


FIGURE 15-18 Image pyramid formed from parking lot image.

An image is formed by convolving the original image with a gaussian kernel, such as that shown in Eq. (15-10).

$$K = \begin{bmatrix} 0.0025 & 0.0125 & 0.0200 & 0.0125 & 0.0025 \\ 0.0125 & 0.0625 & 0.1000 & 0.0625 & 0.0125 \\ 0.0200 & 0.1000 & 0.1600 & 0.1000 & 0.0200 \\ 0.0125 & 0.0625 & 0.1000 & 0.0625 & 0.0125 \\ 0.0025 & 0.0125 & 0.0200 & 0.0125 & 0.0025 \end{bmatrix} \tag{15-10}$$

Notice that the weights in this kernel are largest in the center and fall off gradually away from the

center. This pattern of weights mimics the behavior of a gaussian distribution in two dimensions.

The convolution is performed in the usual fashion, with the exception that instead of the moving window shifting one row or column at a time, it is shifted two rows or columns. This results in a convolved image having one-half as many rows and columns as the original. This reduction in number of pixels is offset by the fact that each pixel in the convolved image represents an area twice the width and height of the original pixels. Each successive convolution is performed on the previously convolved image, resulting in the series of half-resolution copies.

Image pyramids can be used for many purposes. One particularly important use is for multiresolution image matching. By matching images in upper layers of the pyramid, the location of the match can be predicted in lower layers within a couple of pixels, which avoids searching through the entire full-resolution image to find a matching feature. Another use is for quick display of an image while zooming in or out. By first constructing an image pyramid, an image zoom operation can be accomplished by accessing different layers of the pyramid. These are just two possible uses for image pyramids, and there are many others.

15-8 Digital Image Matching

A common task associated with the use of stereoplotters is to place the floating mark at a three-dimensional model position of an object point. This requires the ability to recognize similar image characteristics (texture, shapes, etc.) in small regions from the left and right images of a stereopair. Figure 15-19 shows the same object (in this case a manhole) as it appears in two different aerial photos. The task is to find, using the patch of pixels shown in the left image, the associated patch of pixels in the right image using some measure of similarity between the two. The human visual system is able to perform this task with little conscious effort. When our eyes fixate on an object, the two images merge and the three-dimensional nature of the object is revealed. When one is working with digital images in a softcopy stereoplotter, placing of the floating mark can be performed either manually or by computer processing. Computer software which accomplishes this task uses techniques of digital image matching.

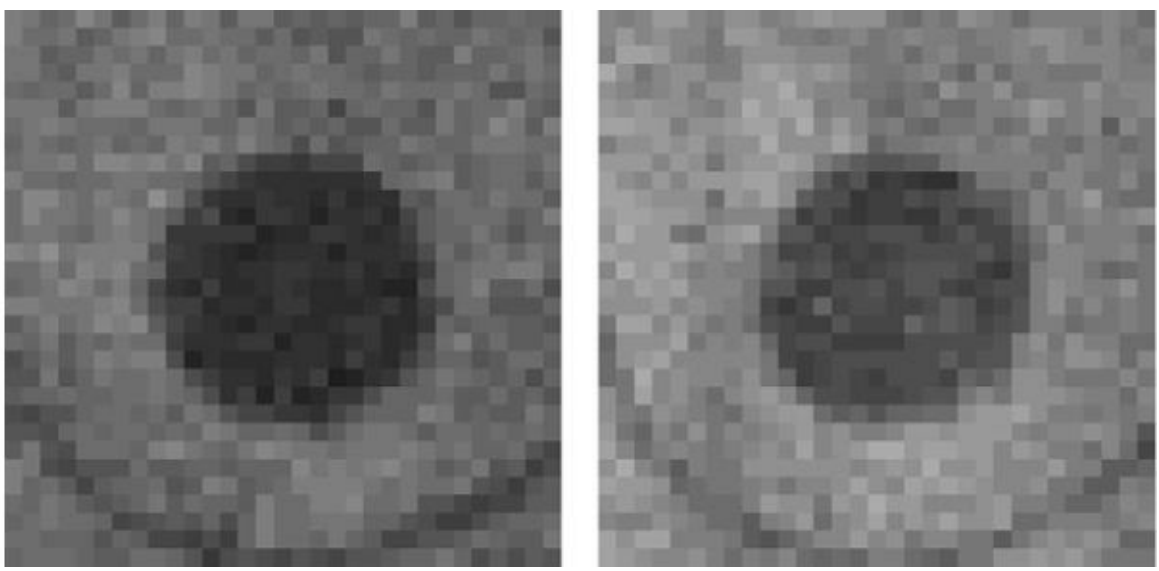


FIGURE 15-19 A manhole in two different aerial photographs.

Digital image-matching techniques fall into three general categories: *area-based*, *feature-based*,

and *hybrid methods*. Area-based methods perform the image match by a numerical comparison of digital numbers in small subarrays from each image. This approach is straightforward and commonly used in softcopy systems. Feature-based methods are more complicated and involve extraction of features, which are comprised of edges at different scales, with subsequent comparison based on feature characteristics such as size and shape. Feature-based image matching requires techniques from the realm of artificial intelligence in computer science. Hybrid methods involve some combination of the first two approaches. Typically, hybrid methods involve preprocessing of the left and right images to highlight features (edges) by methods which were introduced in Secs. 15-6 and 15-7. After the features have been located, they are matched by area-based methods. While all three approaches have particular advantages and disadvantages, this section focuses on area-based image-matching techniques.

Perhaps the simplest area-based digital image-matching method is a technique known as *normalized cross-correlation*. In this approach, a statistical comparison is computed from digital numbers taken from same-size subarrays in the left and right images. A correlation coefficient is computed by the following equation, using digital numbers from subarrays *A* and *B*.

$$C = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}[(A_{ij}-\bar{A})(B_{ij}-\bar{B})]}{\sqrt{\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij}-\bar{A})^2\right]\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(B_{ij}-\bar{B})^2\right]}} \tag{15-11}$$

In Eq. (15-11), c is the correlation coefficient; m and n are the numbers of rows and columns, respectively, in the subarrays; A_{ij} is the digital number from subarray A at row i, column j; $\bar{A}$ is the average of all digital numbers in subarray A; B_{ij} is the digital number from subarray B at row i, column j; $\bar{B}$ and is the average of all digital numbers in subarray B. The correlation coefficient can range from –1 to +1, with +1 indicating perfect correlation (an exact match). A coefficient of –1 indicates negative correlation, which would occur if identical images from a photographic negative and positive were being compared. Coefficient values near zero indicate a nonmatch, and could result from a comparison of any two sets of random numbers. Due to factors such as image noise, perfect (+1) correlation is extremely rare. Generally a threshold value, such as 0.7, is chosen and if the correlation coefficient exceeds that value, the subarrays are assumed to match.

Normalized cross-correlation is essentially the same operation as *linear regression* in statistics. Details of linear regression can be found in most elementary statistics texts, and only general concepts are discussed here. In linear regression, a set of ordered pairs (abscissas and ordinates) is statistically analyzed to determine how well the numbers correspond to a straight-line relationship. In the process, most-probable values are determined for the parameters (slope and intercept) of a best-fit line through the data points. For example, assume the following pair of 3 × 3 arrays of digital numbers are to be analyzed using linear regression.

$$A = \begin{bmatrix} 25 & 48 & 89 \\ 43 & 94 & 47 \\ 76 & 21 & 57 \end{bmatrix} \qquad B = \begin{bmatrix} 33 & 56 & 81 \\ 40 & 98 & 54 \\ 84 & 16 & 49 \end{bmatrix}$$

To compute the linear regression, a tabular solution can be used, as shown in Table 15-1. In this table, the abscissas and ordinates used for the regression are listed in the columns labeled *a* and *b*,

respectively; a^2 and b^2 are their corresponding squares; and $a \times b$ are the products. (*Note:* In typical notation used for linear regression, x_i and y_i are used for abscissas and ordinates, respectively, and this notation is also shown in the table.)

$a(x_i)$	$b(y_i)$	$a^2(x_i^2)$	$b^2(y_i^2)$	$a \times b\ (x_i y_i)$
25	33	625	1089	825
48	56	2304	3136	2688
89	81	7921	6561	7209
43	40	1849	1600	1720
94	98	8836	9604	9212
47	54	2209	2916	2538
76	84	5776	7056	6384
21	16	441	256	336
57	49	3249	2401	2793
$\Sigma = 500$	$\Sigma = 511$	$\Sigma = 33{,}210$	$\Sigma = 34{,}619$	$\Sigma = 33{,}705$

TABLE 15-1 Tabular Form for Computing Linear Regression

In order to compute the regression, the following terms are computed.

$$S_x^2 = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n} = 33{,}210 - \frac{500^2}{9} = 5432.2 \tag{15-12}$$

$$S_y^2 = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n} = 34{,}619 - \frac{511^2}{9} = 5605.6 \tag{15-13}$$

$$S_{xy} = \sum [(x_i - \bar{x})(y_1 - \bar{y})] = \sum (x_i y_i) - \frac{\left(\sum x_i\right)\left(\sum y_i\right)}{n}$$

$$= 33{,}705 - \frac{500 \times 511}{9} = 5316.1 \tag{15-14}$$

$$\beta = \frac{S_{xy}}{S_x^2} = \frac{5316.1}{5432.2} = 0.979 \tag{15-15}$$

$$\alpha = \frac{\sum y_i}{n} - \beta \frac{\sum x_i}{n} = \frac{511}{9} - 0.979\left(\frac{500}{9}\right) = 2.41 \tag{15-16}$$

$$r = \frac{s_{xy}}{\sqrt{S_x^2 S_y^2}} = \frac{33{,}705}{\sqrt{33{,}210 \times 34{,}619}} = 0.963 \tag{15-17}$$

In Eqs. (15-12) through (15-17), n is the number of data points (9 in this case), and the other terms are as indicated in the table. Parameter β in Eq. (15-15) is the slope of the regression line; parameter α in Eq. (15-16) is the y intercept of the regression line; and parameter r in Eq. (15-17) is the *sample correlation coefficient*. Note that r is the same as the normalized cross-correlation coefficient c from Eq. (15-11). Figure 15-17 shows a plot of the nine data points, along with the regression line. In this figure, the nine data points lie nearly along the regression line, which is also indicated by the correlation coefficient r being nearly equal to 1.

Digital image matching by correlation can be performed in the following manner. A candidate subarray from the left photo is chosen, and a search will be performed for its corresponding subarray in the right image. Since the exact position of the image in the right image is not initially known, a search array is selected with dimensions much larger than those of the candidate subarray. A moving window approach is then used, comparing the candidate subarray from the left image with all possible window locations within the search array from the right image, as illustrated in Fig. 15-21. At each window location in the search array, the correlation coefficient is computed in a manner similar to moving window convolution (see Sec. 15-6), resulting in a correlation matrix C. After all coefficients have been calculated, the largest correlation value in C is tested to see if it is above the threshold. If it exceeds the threshold, the corresponding location within the search array is considered to be the match.

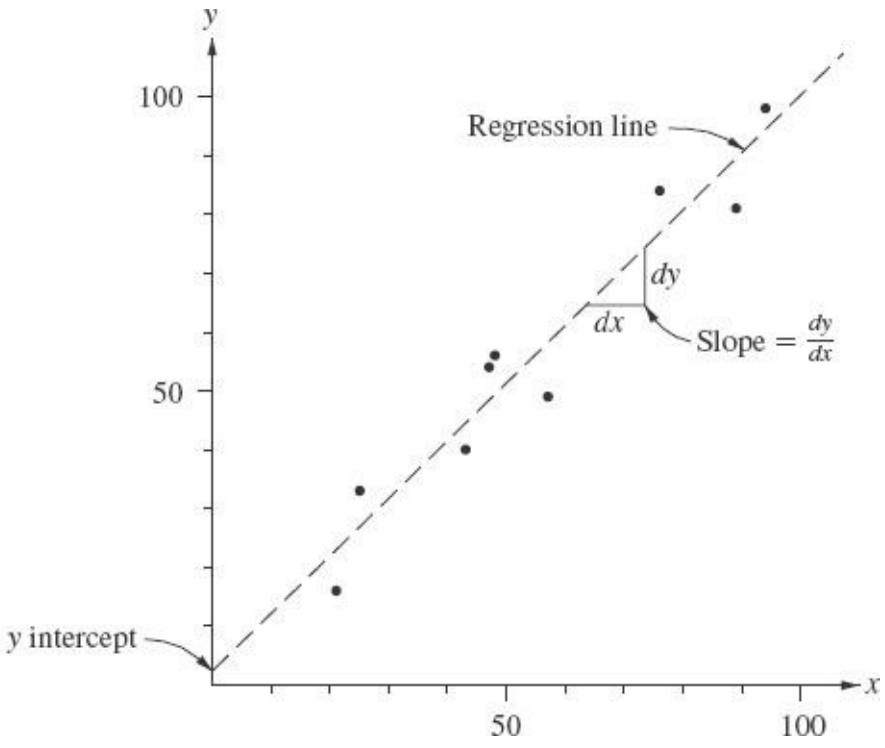


FIGURE 15-20 Linear regression.

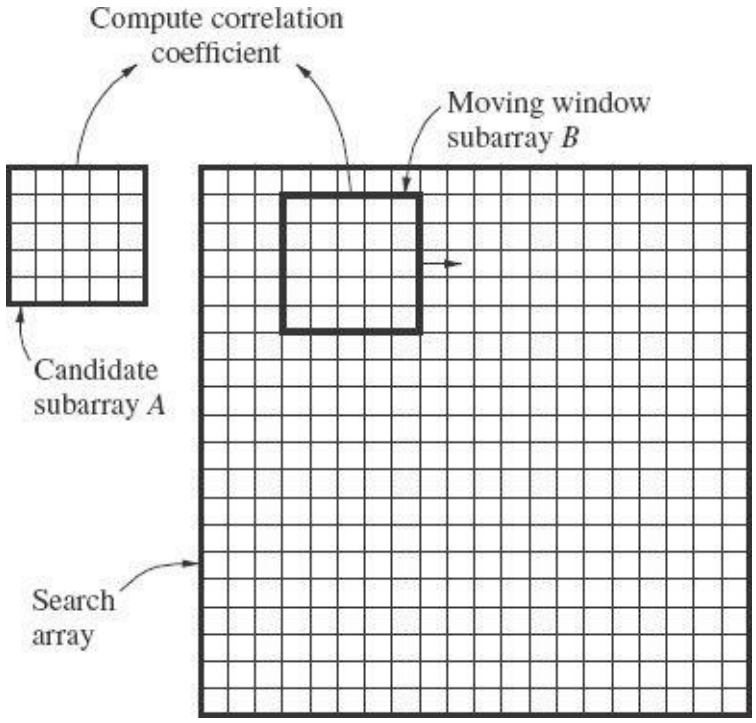


FIGURE 15-21 Computing correlation coefficients using a moving window within a search array.

Example 15-2

The candidate array *A* is an ideal template for a fiducial cross, and the following search array *S* is a portion of a digital image containing a fiducial cross. Compute the position of the fiducial within the search array by correlation.

$$A = \begin{bmatrix} 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 50 & 0 & 0 \\ 50 & 50 & 50 & 50 & 50 \\ 0 & 0 & 50 & 0 & 0 \\ 0 & 0 & 50 & 0 & 0 \end{bmatrix} \quad S = \begin{bmatrix} 41 & 43 & 43 & 49 & 60 & 43 & 41 & 40 & 44 \\ 43 & 44 & 45 & 50 & 64 & 45 & 43 & 43 & 45 \\ 42 & 43 & 44 & 48 & 63 & 49 & 45 & 42 & 42 \\ 42 & 45 & 47 & 50 & 65 & 45 & 45 & 41 & 41 \\ 59 & 62 & 62 & 64 & 69 & 64 & 62 & 63 & 60 \\ 50 & 48 & 48 & 51 & 68 & 55 & 50 & 54 & 53 \\ 42 & 41 & 44 & 48 & 63 & 42 & 47 & 47 & 45 \\ 42 & 44 & 42 & 45 & 62 & 44 & 44 & 45 & 43 \\ 42 & 43 & 44 & 48 & 60 & 47 & 44 & 38 & 35 \end{bmatrix}$$

Solution

Note: The correlation coefficient at the first window position (with the upper left element of a 5 × 5 subarray at the 1, 1 position of the search array) will be calculated as an example.

1. Extract subarray *B* from the search array at position 1, 1.

$$B = \begin{bmatrix} 41 & 43 & 43 & 49 & 60 \\ 43 & 44 & 45 & 50 & 64 \\ 42 & 43 & 44 & 48 & 63 \\ 42 & 45 & 47 & 50 & 65 \\ 59 & 62 & 62 & 64 & 69 \end{bmatrix}$$

2. Compute the average digital number for subarrays *A* and *B*.

$$\bar{A} = \frac{0+0+50+\cdots+50+0+0}{25} = 18$$

$$\bar{B} = \frac{41+43+43+\cdots+62+64+69}{25} = 51.48$$

3. Compute the summation terms for the correlation coefficient.

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n}[(A_{ij}-\bar{A})(B_{ij}-\bar{B})] &= (0-18)(41-51.48)\\ &\quad +(0-18)(43-51.48)\\ &\quad +(50-18)(43-51.48)+\cdots+(50-18)\\ &\quad (62-51.48)+(0-18)(64-51.48)\\ &\quad +(0-18)(69-51.48)\\ &= -1316\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij}-\bar{A})^2 &= (0-18)^2+(0-18)^2+(50-18)^2+\cdots+(50-18)^2\\ &\quad +(0-18)^2+(0-18)^2\\ &= 14,400\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n}(B_{ij}-\bar{B})^2 &= (41-51.48)^2+(43-51.48)^2+(43-51.48)^2\\ &\quad +\cdots+(62-51.48)^2+(64-51.48)^2+(69-51.48)^2\\ &= 2102.24\end{aligned}$$

4. Compute the correlation coefficient.

$$c_{11} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}[(A_{ij}-\bar{A})(B_{ij}-\bar{B})]}{\sqrt{\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij}-\bar{A})^2\right]\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(B_{ij}-\bar{B})^2\right]}} = \frac{-1316}{\sqrt{2102.24\times 14,400}} = -0.24$$

5. Compute the remaining coefficients in the same manner.

$$C = \begin{bmatrix} -0.24 & -0.09 & 0.35 & -0.19 & -0.19 \\ -0.24 & -0.16 & 0.32 & -0.21 & -0.32 \\ 0.25 & 0.37 & 0.94 & 0.29 & 0.27 \\ -0.08 & 0.02 & 0.50 & -0.06 & 0.03 \\ -0.28 & -0.23 & 0.27 & -0.18 & -0.22 \end{bmatrix}$$

6. Select the maximum correlation coefficient. The maximum value, 0.94, occurs at row 3,

column 3 of the *C* array. This value was computed when the upper left element of the moving window was at that position (row = 3, column = 3) in the search array. Since the center of the cross in the template is 2 columns to the right and 2 rows down from the upper left corner, the center of the cross in the search array is at row 3 + 2 = 5 and column 3 + 2 = 5. ▲

A second area-based digital image-matching method is the *least squares matching* technique. (The reader may wish to refer to App. B for general information on least squares.) Conceptually, least squares matching is closely related to the correlation method, with the added advantage of being able to obtain the match location to a fraction of a pixel. Least squares matching can also account for some distortions caused by perspective differences and rotation between images. Figure 15-22 shows: (*a*) a calibration target, (*b*) a tilted image of a target, and (*c*) the resampled result from image (*b*), which is a least squares match of the template (a). Different implementations of least squares matching have been devised, with the following form being commonly used.

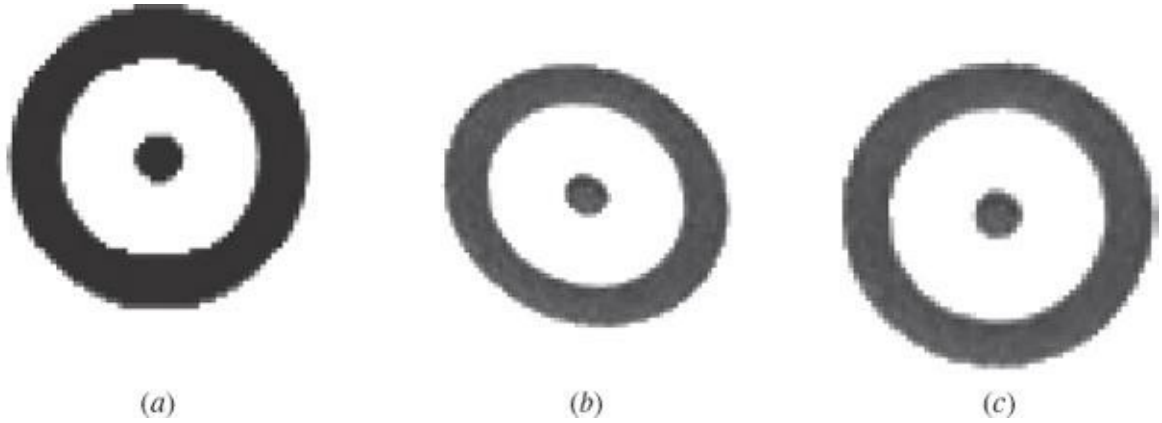


FIGURE 15-22 A calibration target template, the image of the target, and its resampled image.

$$A(x, y) = h_0 + h_1 B(x', y') \tag{15-18}$$

$$x' = a_0 + a_1 x + a_2 y \tag{15-19}$$

$$y' = b_0 + b_1 x + b_2 y \tag{15-20}$$

In Eq. (15-18), $A(x, y)$ is the digital number from the candidate subarray of the left image at location x, y; $B(x', y')$ is the digital number from a subarray in the search area of the right image at location x', y'; h_0 is the radiometric shift; and h_1 is the radiometric scale. Note that parameter h_0 is the same as the y intercept a of Eq. (15-16), and h_1 is the same as the slope b of Eq. (15-15). Equations (15-19) and (15-20) specify an affine relationship (see Sec. C-6) between the coordinates of the pixel on the left photo and the coordinates of the corresponding pixel on the right photo. Figure 15-23 illustrates the positions of subarrays *A* and *B* in the left and right images. In this figure, the x and y axes are the basis for coordinates on the left image, and the x' and y' axes are the basis for coordinates on the right image. Coordinates in both images are expressed in units of pixels.

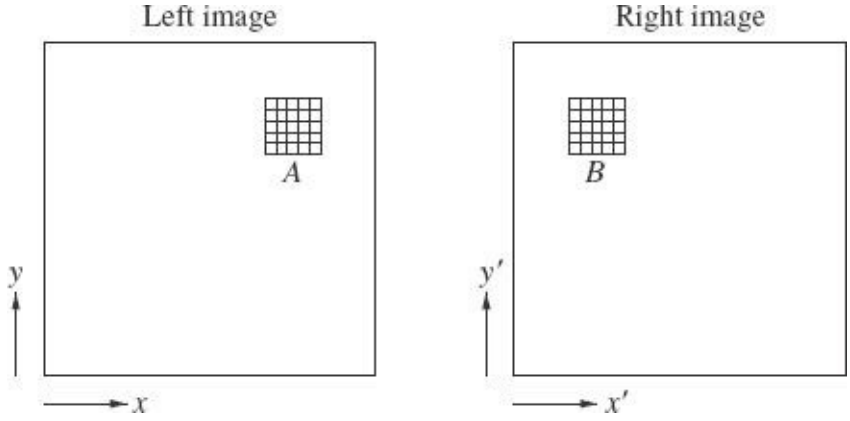


FIGURE 15-23 Positions of subarrays for least squares matching.

Combining Eqs. (15-18), (15-19), and (15-20) and expressing the result in the form of a least squares observation equation gives the following:

$$f = h_0 + h_1 B(a_0 + a_1 x + a_2 y, b_0 + b_1 x + b_2 y) = A(x, y) + V_A \tag{15-21}$$

In Eq. (15-21), f is the function relating the digital numbers from the two images, V_A is the residual, and the other variables are as previously defined. Equation (15-21) is nonlinear, and application of Taylor's series gives the following linear form.

$$\begin{aligned} f = f_0 + f'_{h_0} dh_0 + f'_{h_1} dh_1 + f'_{a_0} da_0 + f'_{a_1} da_1 + f'_{a_2} da_2 + f'_{b_0} db_0 \\ + f'_{b_1} db_1 + f'_{b_2} db_2 = A(x, y) + V_A \end{aligned} \tag{15-22}$$

where

$$\begin{aligned} f'_{h_0} &= 1 \\ f'_{h_1} &= B(x' = , y' =) \\ f'_{a_0} &= h_1 B'_x \\ f'_{a_1} &= x h_1 B'_x \\ f'_{a_2} &= y h_1 B'_x \\ f'_{b_0} &= h_1 B'_y \\ f'_{b_1} &= x h_1 B'_y \\ f'_{b_2} &= y h_1 B'_y \end{aligned}$$

and

$$B'_x = \frac{B(x'+1, y') - B(x'-1, y')}{2} \tag{15-23}$$

$$B'_y = \frac{B(x', y'+1) - B(x'-, y'-1)}{2} \tag{15-24}$$

Since the function f of Eq. (15-21) includes digital numbers from subarray B, partial derivative terms must be obtained using discrete values to estimate the slope of B in both the x and y directions. Equation (15-23) computes the estimate for slope in the x direction by taking the difference between the digital numbers of pixels to the right and left divided by 2, and Eq. (15-24) computes the estimate for slope in the y direction in a corresponding manner. Use of these discrete slope estimates, in conjunction with the chain rule from calculus, allows the partial derivatives (the f' terms) to be determined, as listed above.

Least squares matching is an iterative process which requires an accurate estimate for the position of B within the right image. Initial approximations must be obtained for the unknown parameters h_0, h_1, a_0, a_1, a_2, b_0, b_1, and b_2. Estimates for h_0 and h_1 can be obtained by linear regression as illustrated earlier in this section. If the coordinates of the lower left pixels of A and B are x_0, y_0 and x'_0, y'_0, respectively, the following initial approximations can be used for the affine parameters.

$$a_o = x'_0 - x_0 \qquad a_1 = 1 \qquad a_2 = 0$$

$$b_o = y'_0 - y_0 \qquad b_1 = 0 \qquad b_2 = 1$$

Each iteration of the solution involves forming the linearized equations, solving the equations by least squares to obtain corrections to the approximations, and adding the corrections to the approximations. At the beginning of an iteration, the pixels of subarray B (along with a 1-pixel-wide border around B which is needed for derivative estimates) are resampled (see App. E) from the right image. This is done by stepping through the pixels of subarray A, taking the x and y coordinates of each pixel, and transforming them to the right image x' and y' by using Eqs. (15-19) and (15-20). A corresponding digital number is then resampled from the right image at position x', y'. Once subarray B has been filled, the least squares equations can be formed and solved. The solution is then iterated until the corrections become negligible. On the final iteration, the resampled subarray should be very similar to the template array as Fig. 15-22*c* is to Fig. 15-22*a*.

Some final comments are appropriate at this point. First, the estimated position of subarray B should be within a couple of pixels in order for the solution to converge properly. This can be achieved efficiently through the use of an image pyramid (see Sec. 15-7). Corresponding points can be matched at an upper level of the pyramid where the search area contains fewer pixels. Once the point is matched at a particular level of the pyramid, the position on the next-lower level will be known to within 2 pixels. By progressively matching from upper levels down to the bottom level, accurate position estimates can be obtained at each subsequent level. Another concern is the size of the subarrays to be matched. Generally, a subarray size of 20 × 20 to 30 × 30 gives satisfactory results. If the subarray is much smaller, the low redundancy can result in a weak solution. Larger subarrays can lead to problems due to terrain variations within the image area causing distortions that are not affine. Finally, the transformation equation for y' can be simplified by performing *epipolar resampling* on the images prior to matching. In epipolar resampling, the images are resampled so that the rows line up with epipolar lines. When this is done, Eqs. (15-19) and (15-20) can be simplified to

$$x' = a_0 + a_1 x \tag{15-25}$$

$$y' = b_0 \tag{15-26}$$

15-9 Summary

The discussion of digital image processing methods presented in this chapter is only a brief treatment of the subject. These methods dealt primarily with edge detection, contrast enhancement, noise removal, multiscale representations, and image matching. A wealth of other digital image processing methods are available for dealing with these problems as well as many others. An in-depth understanding of digital image processing requires a great deal of study and experimentation. Those who are interested in learning more about this subject are directed to the references listed below. Further information can be gathered by consulting the bibliographies contained in these documents.

References

Agouris, P., and T. Schenk: "Automated Aerotriangulation Using Multiple Image Multipoint Matching," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 6, 1996, p. 703.

American Society of Photogrammetry: *Manual of Remote Sensing*, vols. 1 and 2, 3d ed., Bethesda, MD, 1998.

American Society for Photogrammetry and Remote Sensing: *Digital Photogrammetry, An Addendum to the Manual of Photogrammetry*, Bethesda, MD, 1996.

———: Special Image Processing Issue, *Photogrammetric Engineering and Remote Sensing*, vol. 55, no. 9, 1989.

———: Special Image Processing Issue, *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 1, 1990.

———: Special Issue on Geostatistics and Scaling of Remote Sensing and Spatial Data, *Photogrammetric Engineering and Remote Sensing*, vol. 65, no. 1, 1999.

Atkinson, P., and P. Curran: "Choosing an Appropriate Spatial Resolution for Remote Sensing Investigations," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 12, 1997, p. 1345.

Baxes, G.: *Digital Image Processing*, Wiley, New York, 1994.

Burt, P. J.: "The Pyramid as a Structure for Efficient Computation," *Multiresolution Image Processing and Analysis*, ed. A. Rosenfeld, Springer-Verlag, Berlin, 1984.

Carnahan, W., and G. Zhou: "Fourier Transform Techniques for the Evaluation of the Thematic Mapper Line Spread Function," *Photogrammetric Engineering and Remote Sensing*, vol. 52, no. 5, 1986, p. 639.

Chavez, P.: "Radiometric Calibration of Landsat Thematic Mapper Multispectral Images," *Photogrammetric Engineering and Remote Sensing*, vol. 55, no. 9, 1989, p. 1285.

———: "Image Based Atmospheric Corrections—Revisited and Improved," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 9, 1996, p. 1025.

Chui, C. K.: *An Introduction to Wavelets*, Academic Press, San Diego, CA, 1992.

Crippen, R.: "A Simple Spatial Filtering Routine for the Cosmetic Removal of Scan-Line Noise from Landsat TM P-Tape Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 55, no. 3, 1989, p. 327.

Ekstrom, M.: *Digital Image Processing Techniques*, Academic Press, New York, 1984.

Ekstrom, M., and A. McEwen: "Adaptive Box Filters for Removal of Random Noise from Digital Images," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 4, 1990, p. 453.

Hannah, M. J.: "A System for Digital Stereo Matching," *Photogrammetric Engineering and Remote Sensing*, vol. 55, no. 12, 1989, p. 1765.

Helava, U. V.: "Object-Space Least-Squares Correlation," *Photogrammetric Engineering and Remote*

Sensing, vol. 54, no. 6, 1988, p. 711.
Hord, R.: *Digital Image Processing of Remotely Sensed Data*, Academic Press, New York, 1982.
Horgan, G.: "Wavelets for SAR Image Smoothing," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 12, 1998, p. 1171.
Hunt, B. R., T. W. Ryan, and F. A. Gifford: "Hough Transform Extraction of Cartographic Calibration Marks from Aerial Photography," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 7, 1993, p. 1161.
Jensen, J.: *Introductory Digital Image Processing: A Remote Sensing Perspective*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ, 2004.
Kardoulas, N., A. C. Bird, and A. I. Lawan: "Geometric Correction of SPOT and Landsat Imagery: A Comparison of Map and GPS Derived Control Points," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 10, 1996, p. 1173.
Kruphik, A.: "Using Theoretical Intensity Values as Unknowns in Multiple-Patch Least-Squares Matching,"*Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 10, 1996, p. 1151.
Li, M.: "Hierarchical Multipoint Matching," *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 8, 1991, p. 1039.
Light, D.: "Film Cameras or Digital Sensors? The Challenge," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 3, 1996, p. 285.
Lillesand, T., and R. Kiefer: *Remote Sensing and Image Interpretation*, 6th ed., Wiley, New York, 2008.
Lim, J. S.: *Two-Dimensional Signal and Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1990.
Moik, J. G.: *Digital Processing of Remotely Sensed Images*, NASA SP-431, Government Printing Office, Washington, 1980.
Muller, J.: *Digital Image Processing in Remote Sensing*, Taylor & Francis, London, 1986.
Novack, K.: "Rectification of Digital Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 3, 1992, p. 339.
Pratt, W. K.: *Digital Image Processing*, 4th ed., Wiley, New York, 2007.
Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery: *Numerical Recipes in C*, Cambridge University Press, New York, 1994.
Richards, J.: *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, New York, 1986.
Schenk, T,: *Digital Photogrammetry*, Laurelvile, OH, 1999.
Stollnitz, E.J., T. DeRose, and D. H. Salesin: "Wavelets for Computer Graphics: A primer, part 1," *IEEE Computer Graphics and Applications*, vol. 15, 1995, p. 76.
Thome, K., B. Markham, J. Barker, P. Slater, and S. Bigger: "Radiometric Calibration of Landsat," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 7, 1997, p. 853.
Tseng, Y., J. Tzen, K. Tang, and S. Lin: "Image-to-Image Registration by Matching Area Features Using Fourier Descriptors and Neural Networks," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 8, 1997, p. 975.
Walsh, S., J. W. Cooper, I. E. Von Essen, and K. R. Gallager: "Image Enhancement of Landsat Thematic Mapper Data and GIS Integration for Evaluation of Resource Characteristics," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 8, 1990, p. 1135.
Wilkinson, B. E., B. A. Dewitt, A. C. Watts, A. H. Mohamed, and M. A. Burgess: "A New Approach for Pass-point Generation from Aerial Video Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 75, no. 12, 2009, p. 1415.

Problems

15-1. Give a brief definition of a digital image model.

15-2. Define spatial frequency.

15-3. What is the Nyquist frequency? Why is it important?

15-4. Briefly describe the linear stretch method of contrast enhancement.

15-5. Give a brief description of the histogram equalization method of contrast enhancement.

15-6. Briefly describe the relationship between the spatial domain and frequency domain for a digital image.

15-7. Give a brief description of the discrete Fourier transform.

15-8. What is the fast Fourier transform and why is it important?

15-9. Briefly describe the wavelet transform.

15-10. Name three uses of the wavelet transform.

15-11. Given the following convolution kernel K and the matrix of digital numbers D from a 5×5 image, compute the convolved image C, using the moving window approach.

$$K = \begin{bmatrix} 0.05 & 0.05 & 0.05 \\ 0.05 & 0.60 & 0.05 \\ 0.05 & 0.05 & 0.05 \end{bmatrix}$$

$$D = \begin{bmatrix} 71 & 54 & 54 & 63 & 49 \\ 45 & 63 & 57 & 66 & 53 \\ 88 & 76 & 63 & 65 & 55 \\ 52 & 44 & 61 & 77 & 57 \\ 75 & 76 & 43 & 73 & 59 \end{bmatrix}$$

15-12. Repeat Prob. 14-9, except use the following convolution kernel K.

$$K = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

15-13. Briefly explain the difference between high-pass and bandpass filters.

15-14. Outline a process by which edges at different scales can be extracted from a digital image.

15-15. Briefly explain how an image pyramid is constructed.

15-16. Briefly describe the three categories of digital image matching.

15-17. Describe how normalized cross-correlation relates to linear regression in image matching.

15-18. Given the following two subarrays from a pair of digital images, compute the normalized cross-correlation coefficient.

$$A = \begin{bmatrix} 45 & 33 & 43 & 79 & 56 \\ 65 & 90 & 65 & 54 & 69 \\ 45 & 88 & 57 & 56 & 73 \\ 56 & 56 & 43 & 65 & 65 \\ 76 & 25 & 43 & 77 & 66 \end{bmatrix}$$

$$B = \begin{bmatrix} 52 & 56 & 52 & 77 & 61 \\ 73 & 94 & 73 & 61 & 78 \\ 42 & 99 & 65 & 67 & 82 \\ 64 & 62 & 49 & 73 & 76 \\ 77 & 30 & 54 & 87 & 75 \end{bmatrix}$$

15-19. Repeat Prob. 15-18, except that the following subarrays are to be used.

$$A = \begin{bmatrix} 35 & 22 & 14 & 34 & 89 \\ 22 & 10 & 17 & 19 & 50 \\ 84 & 10 & 51 & 95 & 31 \\ 97 & 46 & 19 & 81 & 65 \\ 15 & 9 & 23 & 19 & 80 \end{bmatrix}$$

$$B = \begin{bmatrix} 44 & 35 & 42 & 98 & 118 \\ 33 & 44 & 11 & 37 & 100 \\ 95 & 23 & 77 & 121 & 66 \\ 91 & 77 & 55 & 120 & 95 \\ 21 & 23 & 79 & 27 & 124 \end{bmatrix}$$

15-20. Discuss the differences between digital image matching by normalized cross-correlation and least squares matching.

CHAPTER 16
Control for Aerial Photogrammetry

16-1 Introduction

Photogrammetric control traditionally consists of any points whose positions are known in an object-space reference coordinate system and whose images can be positively identified in the photographs. In aerial photogrammetry the object space can be defined by various reference ground coordinate systems (see Chap. 5). Photogrammetric *ground control*, as it is commonly called in aerial photogrammetry, provides the means for orienting or relating aerial photographs to the ground. Almost every phase of photogrammetric work requires some ground control. Aircraft-mounted GPS/INS systems can also be used as control for photogrammetric surveys by directly measuring the position and attitude of the camera during acquisition.

Photogrammetric ground control is generally classified as either *horizontal control* (the position of the point in object space is known with respect to a horizontal datum) or *vertical control* (the elevation of the point is known with respect to a vertical datum). Separate classifications of horizontal and vertical control have resulted primarily because of differences in horizontal and vertical reference datums, and because of differences in surveying techniques for establishing horizontal and vertical control. Also, horizontal and vertical control are considered separately in some photogrammetric processes. Often, however, both horizontal and vertical object-space positions of points are known, so that these points serve a dual control purpose.

Field surveying for photogrammetric control has historically been a two-step process, although now with the widespread use of GPS, this distinction is not as clear. The first step consists of establishing a network of *basic control* in the project area. This basic control consists of horizontal control monuments and benchmarks of vertical control which will serve as a reference framework for subsequent photo control surveys. The second step involves establishing object-space positions of *photo control* by means of surveys originating from the basic control network. Photo control points are the actual image points appearing in the photos that are used to control photogrammetric operations. The accuracy of basic control surveys is generally higher than that of subsequent photo control surveys. If GPS is used for the control surveying work, in some cases the intermediate step of establishing basic control may be bypassed and photo control established directly. This is discussed further in Sec. 16-5.

The two-step procedure of field surveying for photogrammetric control is illustrated in Fig. 16-1. In the figure a basic GPS control survey originates from existing control stations E_1 through E_4 and establishes a network of basic control points B_1 through B_6 in the project area. With these basic stations established, the second step of conducting subordinate surveys to locate photo control can occur. This is illustrated with the surveys that run between B_5 and B_6 and locate photo control points P_1 and P_2.

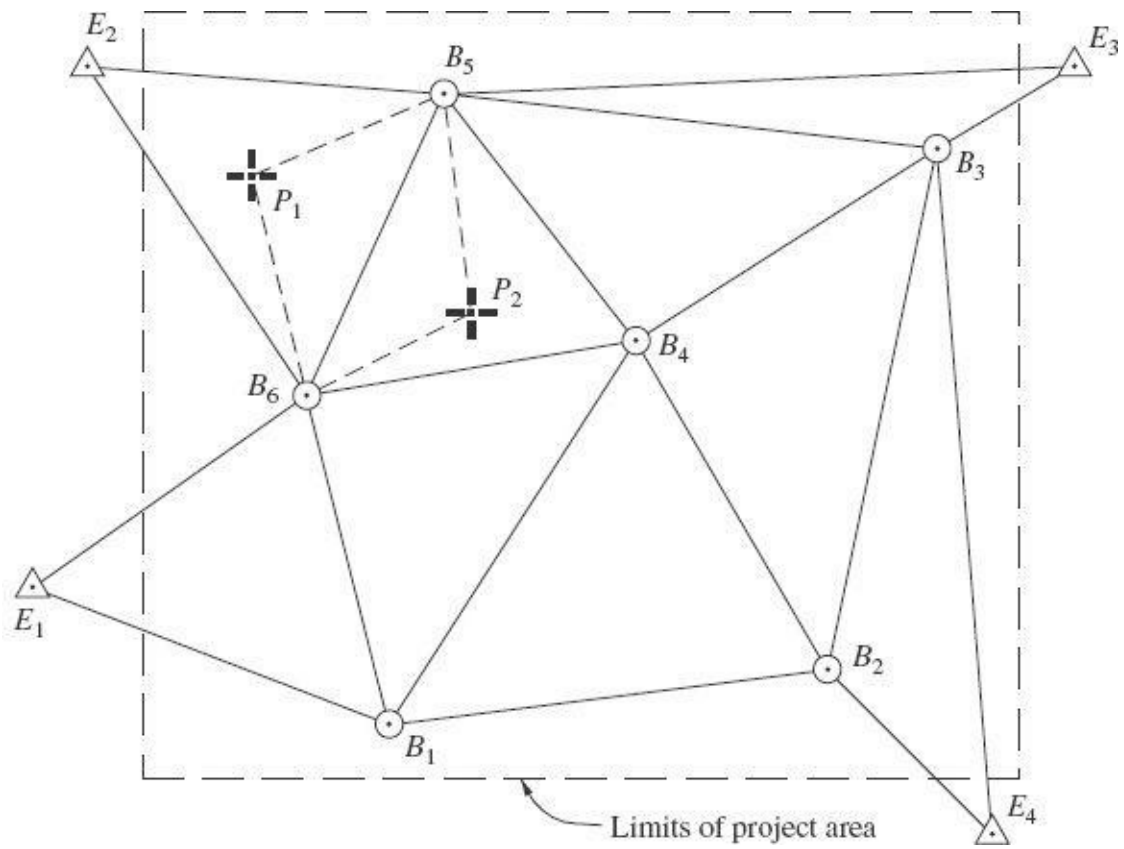


FIGURE 16-1 Field surveys for establishing photogrammetric control.

The establishment of good ground control is an extremely important element of almost any type of photogrammetric project. The accuracy of finished photogrammetric products can be no better than the ground control upon which they are based. Many maps and other items that have been carefully prepared in the office to exacting standards have failed to pass field inspection simply because the ground control was of poor quality. Because of its importance, the ground control phase of every photogrammetric project should be carefully planned and executed. The cost of establishing ground control for photogrammetric mapping is normally substantial, and depending upon conditions it can be expected to constitute between 10 and 50 percent of the total project cost.

Planning is an essential step before carrying out a photogrammetric ground control survey. The required accuracy and consequently, the type of equipment and field techniques to be used should be settled early. Required accuracy of photo control depends primarily upon the accuracy necessary in the photogrammetric mapping or other project that it controls. However, another consideration is whether the control will serve other purposes in addition to controlling the photogrammetric work. For example, if a map is to be prepared for planning and designing an urban transportation system, the ground control should be accurate enough to also be used in describing property right of way acquisition and to serve in properly laying out the construction alignment.

The *national map accuracy standards* (NMAS) are commonly used to govern accuracy requirements of maps and therefore indirectly will govern control surveying accuracy. If a map produced photo-grammetrically is to be labeled as complying with these standards, it is required that at least 90 percent of the principal planimetric features be plotted to within 1/30 in (0.8 mm) of their true positions for map scales of 1:20,000 or larger, and to within 1/50 in (0.5 mm) for scales smaller than 1:20,000. On a map plotted at a scale of 1:600 (1 in/50 ft), this represents an allowable horizontal map error of 0.5 m (1.7 ft) on the ground. On a map at a scale of 1:24,000 (1 in/2000 ft), the allowable horizontal map error is 12 m (40 ft). If NMAS are to be met, horizontal photo control must be located to greater accuracy than the allowable horizontal map error. As a general rule of thumb, photo control

should contain error no greater than about one-fourth to one-third the horizontal map accuracy tolerance. Some organizations require stricter tolerances than this. Of course, basic control must be more accurate than photo control.

NMAS also require that at least 90 percent of all points tested for elevation be correct to within one-half the contour interval. To meet this standard, a rule of thumb in topographic mapping states that elevations of vertical photo control points should be correct to within plus or minus about one-fifth of the contour interval; but as an additional safety factor, some agencies require that their accuracy be within one-tenth of the contour interval. According to this latter rule, a map being plotted with a contour interval of 1 m requires vertical photo control accurate to within 0.1 m. Again, the basic control must be more accurate than this.

A more current set of accuracy standards drafted by the Federal Geographic Data Committee is titled *Geospatial Positioning Accuracy Standards*. These standards are readily applicable to digital maps, which are stored in a computer and manipulated with CAD software. As with NMAS, accuracy is expressed with separate horizontal and vertical components. Other sets of accuracy standards have also been established by various organizations such as the American Society for Photogrammetry and Remote Sensing, the Federal Highway Administration, and the American Society of Civil Engineers.

In planning the control survey, maximum advantage should be taken of existing control in the area. The National Geodetic Survey has established numerous horizontal control monuments and vertical control benchmarks in its work of extending the national control network. The U.S. Geological Survey has also established a network of reliable horizontal and vertical control monuments in its topographic mapping operations. In certain localities, other agencies of the federal government such as the Tennessee Valley Authority, Army Corps of Engineers, and Bureau of Land Management have established control. In addition, various state, county, and municipal agencies may have established control monuments.

16-2 Ground Control Images and Artificial Targets

In general, images of acceptable photo control points must satisfy two requirements: (1) They must be sharp, well defined, and positively identified on all photos; and (2) they must lie in favorable locations in the photographs. (Reasons for this latter requirement are described in Sec. 16-3.) Control surveys for photogrammetry are normally conducted after the photography has been obtained. This ensures that the above two requirements can be met. Photo control images are selected after careful study of the photos. The study should include the use of a stereoscope to ensure a clear stereoscopic view of all points selected. This is important because many of the subsequent photogrammetric measurements will be made stereoscopically. A preliminary selection of photo control images may be made in the office, but the final selection should be made in the field with the photos in hand. This enables positive identification of objects to be made, and it also permits making a firsthand assessment of object point accessibility, terrain conditions, and surveying convenience.

When selecting the position of a control point associated with an object, one should also take into consideration the precise measurement of the point in the images that will take place. For example, if the chevron-shaped feature shown in Fig. 16-2*a* is to be used as a control point, it may be tempting to use the point formed by the edges of the two legs. However, due to blurring along the edges, it will be difficult to precisely measure the position of the point in the image. A better choice is to use the intersection of the centerline of the legs, illustrated in Fig. 16-2*b*, which is easier to identify than the edges in the image.

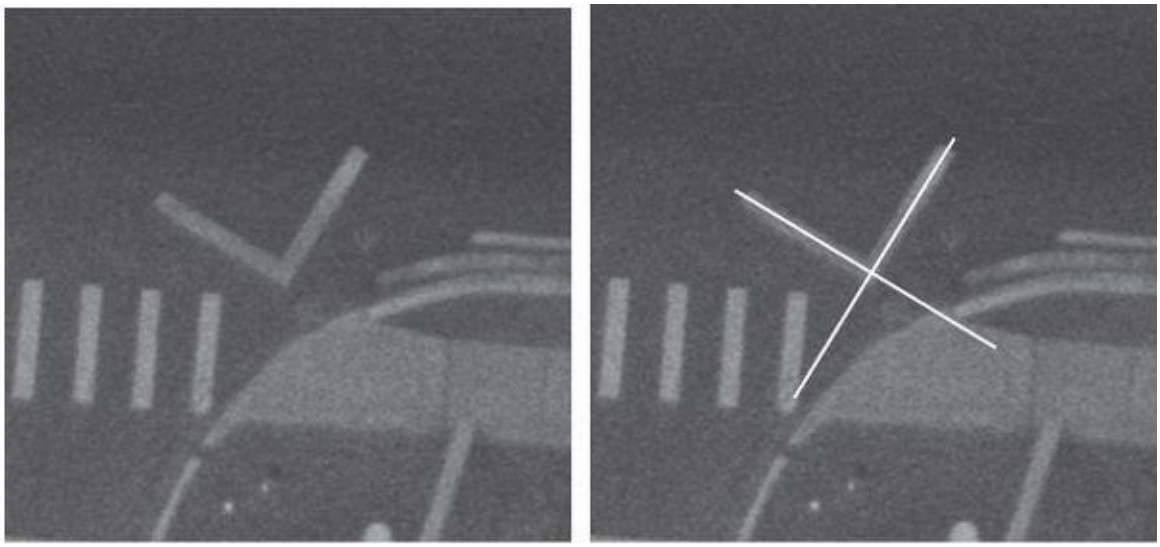


FIGURE 16-2 (*a*) Control point image with blurred edges. (*b*) Intersection of centerlines of legs.

Images for horizontal control have slightly different requirements than images for vertical control. Because their horizontal positions on the photographs must be precisely measured, images of horizontal control points must be very sharp and well defined horizontally. Some objects whose images are commonly satisfactory for horizontal control are intersections of sidewalks, intersections of roads, manhole covers, small lone bushes, isolated rocks, corners of buildings, fence corners, power poles, points on bridges, intersections of small trails or watercourses, etc. Care must be exercised to ensure that control points do not fall in shadowed areas on some photos.

Images for vertical control need not be as sharp and well defined horizontally. Points selected should, however, be well defined vertically. Best vertical control points are small, flat or slightly crowned areas. The small areas should have some natural features nearby, such as trees or rocks, which help to strengthen stereoscopic depth perception. Large, open areas such as the tops of grassy hills or open fields should be avoided, if possible, because of the difficulties they cause in stereoscopic depth perception. Intersections of roads and sidewalks, small patches of grass, small bare spots, etc., make excellent vertical control points.

In some areas such as prairies, forests, and deserts, natural points suitable for photogrammetric control may not exist. In these cases artificial points called *panel points* may be placed on the ground prior to taking the aerial photography. Their positions are then determined by field survey or in some cases by aerotriangulation. This procedure is called *premarking* or *paneling*. Artificial targets provide the best possible photographic images, and therefore they are used for controlling the most precise photogrammetric work, whether or not natural points exist. Artificial targets are also used to mark section corners and boundary lines for photogrammetric cadastral work.

Besides their advantage of excellent image quality, their unique appearance makes misidentification of artificial targets unlikely. Disadvantages of artificial targets are that extra work and expense are incurred in placing the targets, the targets could be moved between the time of their placement and the time of photography, and the targets may not appear in favorable locations on the photographs. To guard against the movement of artificial targets, the photography should be obtained as near as possible to the time of placing targets. To obtain target images in favorable positions on the photographs, the coverage of each photo can be planned in relation to target locations, and the positions of ground principal points can be specified on the flight plan.

A number of different types of artificial targets have been successfully used for photogrammetric

control. The main elements in target design are good color contrast, a symmetric target that can be centered over the control point, and a target size that yields a satisfactory image on the resulting photographs. Contrast is best obtained using light-colored targets against a dark background or dark-colored targets against light backgrounds. The target shown in Fig. 16-3 provides good symmetry for centering over the control point. The middle panel of the target should be centered over the control point, since this is the image point to which measurements will be taken. The legs help in identifying targets on the photos, and also help in determining the exact center of the target should the image of the center panel be unclear. While the target shown in Fig. 16-3 is perhaps the ideal shape, circumstances may dictate use of other target shapes. Figure 16-4*a* shows a target which is often used where a smaller target is needed. The target of Fig. 16-4*b* is nearly as effective as that of Fig. 16-3, and it has the advantage of being more easily and quickly constructed. The target of Fig. 16-4*c* is less than optimal due to lack of biaxial symmetry; however, it may be needed in confined areas such as edges of highways.

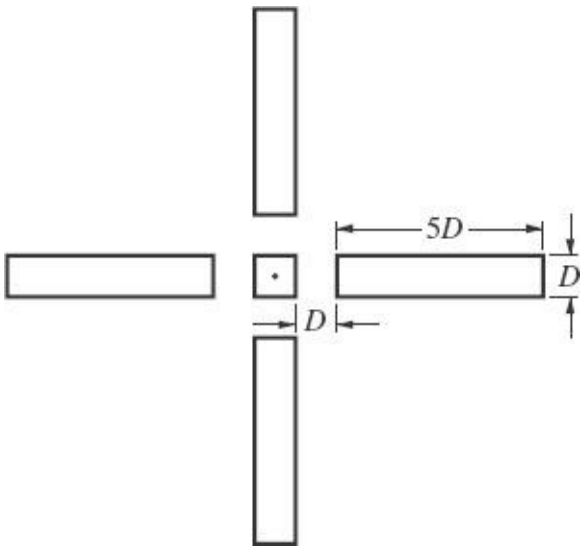


FIGURE 16-3 Artificial photogrammetric target.

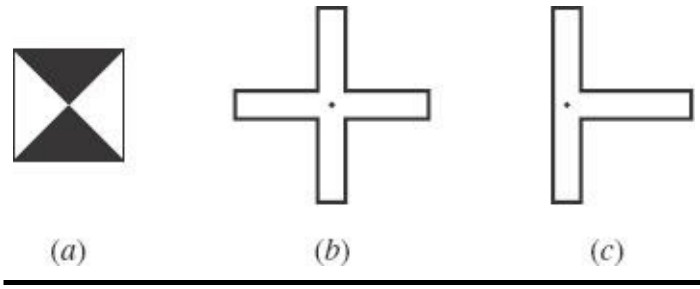


FIGURE 16-4 Other common artificial targets.

Target sizes must be designed on the basis of intended photo scale so that the target images are the desired size on the photos. An image size of about 0.03 mm to about 0.10 mm in film photography and about 10 to 20 pixels in digital photography for the sides of the central panel is suitable. As shown in Fig. 16-3, if the ground dimension of the central panel of the target is *D*, then the leg width should also be *D*, leg length should be 5*D*, and the open space between the central panel and the leg should be *D*. Target sizes are readily calculated once the photo scale and optimum target image size are selected. If, for example, a central panel size of 0.05 mm is desired and photography at a scale of 1:12,000 is planned, then *D* should be 0.6 m.

Materials used for targeting are quite variable. In some cases, satisfactory targets are obtained by simply painting white crosses on blacktop roads. In other cases, targets are painted on plywood, masonite, or heavy cloth, in which case they may be salvaged and reused. Satisfactory targets have also been made by placing stones against an earth background in the shape of a cross. The stones may be painted white for added contrast. Lime placed in the shape of a cross against a dark background has also produced satisfactory targets. Old tires painted white centered over the control points are also good for low-altitude, large-scale photography.

If panel points are needed in an area after the photography has already been obtained a procedure known as *post marking* can be performed. In this method, targets as described above are placed in the desired positions. Supplemental vertical photographs are taken of each target and its surrounding area with a small-format camera carried in a light aircraft flying at low altitude. Flying height can be calculated for the supplemental photography so that it has the same scale as the original photography. Locations of the targets can then be transferred by superimposing scaled supplemental photography over the original photography. A similar method can be performed in film photography using a point transfer device.

The importance of exercising extreme caution in locating and marking objects in the field that correspond to selected photo images cannot be overemphasized. Mistakes in point identification are common and costly. A power pole, for example, may be located in the field, but it may not be the same pole whose image was identified on the photos. Mistakes such as this can be avoided by identifying enough other details in the immediate vicinity of each point so that verification is certain. A pocket stereoscope taken into the field can be invaluable in point identification, not only because it magnifies images but also because hills and valleys which aid in object verification can be seen both on the photos and on the ground.

16-3 Number and Location of Photo Control

The required number of control points and their optimum location in the photos depend upon the use that will be made of them. For a very simple problem such as calculating the flying height of a photo which is assumed to be vertical (see Sec. 6-9), only the horizontal length of a line and the elevations of its endpoints are needed. A line of as great a length as possible should be chosen. For controlling mosaics (see Chap. 9), only a sparse network of horizontal control may be needed. The network should be uniformly distributed throughout the block of photos.

In solving the *space resection* problem for determining the position and orientation of a tilted photo (see Sec. 11-6), a minimum of three *XYZ* control points is required. The images of the control points should ideally form a large, nearly equilateral triangle. Although three control points are the required minimum for space resection, redundant control is recommended to increase the accuracy of the photogrammetric solution and to prevent mistakes from going undetected.

If photo control is being established for the purpose of orienting stereomodels in a plotting instrument for topographic map compilation, the absolute minimum amount of control needed in each stereomodel is three vertical and two horizontal control points. Again, the prudent photogrammetrist will utilize some amount of redundant control. As a practical minimum, each stereomodel oriented in a plotter should have three horizontal and four vertical control points. The horizontal points should be fairly widely spaced, and the vertical control points should be near the corners of the model. A satisfactory configuration is shown in Fig. 16-5. Some organizations require a fifth vertical control point in the center of each stereomodel.

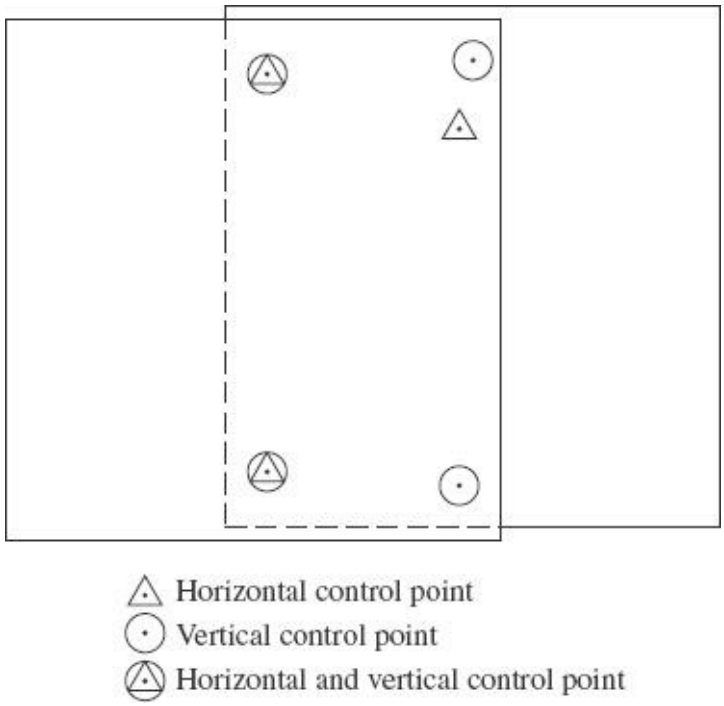
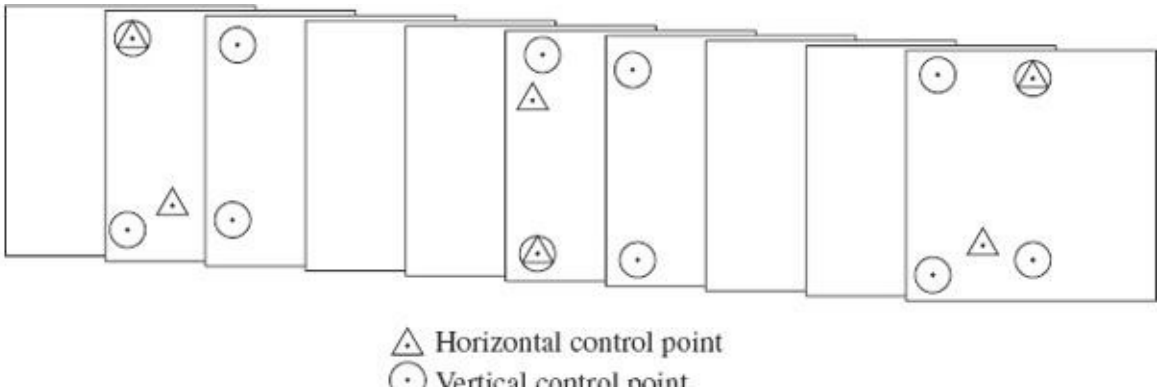


FIGURE 16-5 Control recommended for orienting stereomodels in a stereoscopic plotting instrument.

If aerotriangulation (see Chap. 17) is planned to supplement photo control, then fewer ground-surveyed photo control points are needed. The amount of ground-surveyed photo control needed for aerotriangulation will vary, depending upon the size, shape, and nature of the area to be covered, the resulting accuracy required, and the procedures, instruments, and personnel to be used. In general, the more dense the ground-surveyed network of photo control, the better the resulting accuracy in the supplemental control determined by aerotriangulation. There is an optimum amount of ground-surveyed photo control, however, which affords maximum economic benefit from aerotriangulation and at the same time maintains a satisfactory standard of accuracy. On average, if aerotriangulation of a strip of photos is to be performed for the purpose of obtaining control for orienting stereomodels in a stereoplotter, a minimum of about two horizontal and three or four vertical ground-surveyed photo control points should appear in approximately every fifth stereomodel along the strip. This configuration is shown in Fig. 16-6. For aerotriangulation of blocks of photos, the ground-surveyed control should be systematically arranged throughout the block. Best control configurations consist of horizontal control along the periphery of the block with a uniform distribution of vertical control throughout the block. Experience generally dictates the best control configurations to use, and organizations involved in aerotriangulation normally develop their own standards which meet accuracy requirements for their particular combination of procedures, instruments, and personnel.



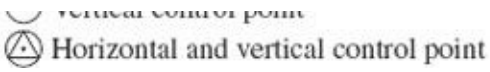


FIGURE 16-6 Example control configuration for a strip of photographs.

16-4 Traditional Field Survey Methods for Establishing Horizontal and Vertical Control

There are numerous and varied traditional instruments and techniques for field surveying. In the context of this chapter, GPS is considered a nontraditional method and is discussed in subsequent sections. This section provides a very brief discussion of some basic methods. The textbooks on surveying listed as references at the end of this chapter provide a much more thorough treatment of these subjects.

Traditional field methods are divided into surveys for horizontal and vertical control. Horizontal control surveys, both for basic control and for photo control, may be conducted using any of the traditional field methods: *traversing*, *triangulation*, or *trilateration*. Total station instruments are most commonly used for measuring the angles and distances associated with these methods. Traversing, the most common of the three mentioned, consists of measuring horizontal angles and horizontal distances between consecutive stations of a closed network. The existing reference control stations are included in the network. Based on the existing control coordinates and reference direction, along with the newly measured horizontal angles and distances, coordinates of all new stations may be calculated trigonometrically in the rectangular coordinate system of the existing reference control. An adjustment, typically least squares (see App. B), is normally made to account for measurement errors.

Triangulation involves measurement of horizontal angles between intervisible stations of a network of triangular figures. Trilateration, on the other hand, involves measurement of horizontal distances in such a network. Often, triangulation and trilateration are combined in a given network. A least squares adjustment can be performed on the combined network to yield coordinates for the new stations. Using triangulation-combined trilateration enables highly accurate coordinates to be determined.

For any horizontal control survey, an important task which must precede the taking of field measurements is establishing the network of stations in the project area whose positions are to be determined. In basic control surveys, the stations will normally be artificial monuments such as wooden stakes or iron rods driven into the ground. These are carefully referenced to permanent nearby features so that they can be recovered later if lost. In photo control surveys, some of the stations will be artificial monuments and others will be the natural features selected for photo control points.

For vertical control surveys, *differential leveling* is a common field procedure. The basic equipment for differential leveling is a leveling instrument and a graduated rod. The leveling instrument consists of a sighting telescope with reticle for reading the graduated rod and a means for orienting the telescope's line of sight in a horizontal plane. Another technique for determining differences in elevation is *trigonometric leveling*. It may be used where moderate accuracy is required and is especially well suited for rugged terrain. In trigonometric leveling, as illustrated in Fig. 16-7, vertical (zenith) angle Z and slope distance S are measured from the instrument at A to the reflector at B. As discussed in the preceding section, these measurements can be conveniently made with a total station instrument. The difference in elevation from the center of the instrument to the center of the reflector is $S \cos Z$. The elevation of point B is then equal to the elevation of point A, plus the instrument height i_h, plus $S \cos Z$, minus the reflector height r_h, as illustrated in Fig. 16-7. Compensation for errors due to earth curvature and atmospheric refraction is generally made by

standard correction formulas found in many surveying texts. For highest accuracy in compensating for these errors, the measurements should be made in both directions and averaged.

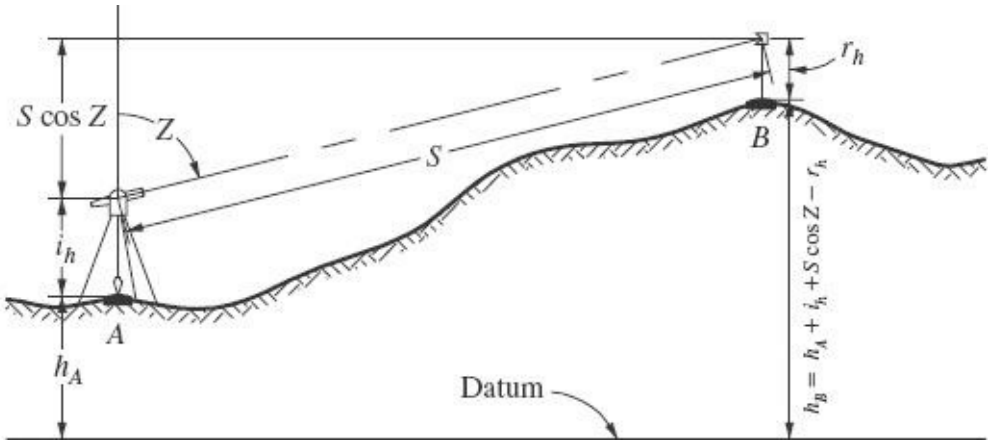


FIGURE 16-7 Trigonometric leveling.

16-5 Fundamentals of the Global Positioning System

GPS was implemented by the U.S. Department of Defense to address the problem of navigation and positioning on a global basis. It currently consists of a network of 31 satellites orbiting the earth at an altitude of approximately 20,000 km, along with associated ground tracking and control stations. The satellites are arranged in six orbital planes and make a complete revolution every 12 hours (h). This arrangement guarantees that at least four satellites will be "visible" from anywhere on earth at any instant in time. The satellites are equipped with transmitters which can broadcast electromagnetic signals, as well as atomic clocks which establish a highly precise and accurate time basis.

Each satellite broadcasts digital messages or codes on two carrier frequencies: L1 at 1575.42 megahertz (MHz) and L2 at 1227.60 MHz. The codes are a pseudorandom series of bits (0s and 1s) which have been modulated onto the carrier frequencies. Two primary codes are broadcast: the C/A or *coarse acquisition* code, which is carried only by the L1 frequency, and the P or *precise* code, which is broadcast on both L1 and L2. In addition to these two primary codes, additional data are broadcast, including timing information and a satellite ephemeris. An additional carrier frequency, L5, is being introduced with the launch of new GPS satellites. The L5 signal is broadcast at 1176.45 MHz, and will add accuracy and robustness to the current system.

The C/A code is broadcast at a *chip rate* of 1.023 MHz or 1.023×10^6 bits per second (bits/s). Given the speed of light of 2.9979246×10^8 m/s, each bit of the C/A code has an *equivalent length* of 293 m. The P code is broadcast at a frequency of 10.23 MHz, giving an equivalent length of 28.3 m/bit. The carrier waves are broadcast at much higher frequencies and therefore have wavelengths which are substantially shorter than the code bits. For the L1 frequency the wavelength is 0.190 m, and for L2 it is 0.244 m. The equivalent lengths of the code bits and the wavelengths of the carrier frequencies relate directly to the precision with which positions can be determined. Thus positions computed through *carrier-phase* processing can be determined much more precisely than those computed through *code-phase* processing.

The primary components of receivers used in conjunction with GPS are an antenna and a module which contains the receiver electronics, a clock, and a computer. Figure 16-8 shows a typical GPS receiver with the antenna set up on a tripod directly over a control monument. Virtually all GPS receivers are capable of measuring the C/A code phase while fewer can measure the P code phase. Only the more expensive *geodetic*-grade receivers can measure the carrier phase. In addition, some

units can receive only the L1 frequency while *dual-frequency* receivers can receive both L1 and L2. Newer receivers have the ability to receive L5 signals, and some can receive signals from global navigation satellite system (GNSS) signals other than GPS, such as the Russian GLONASS.

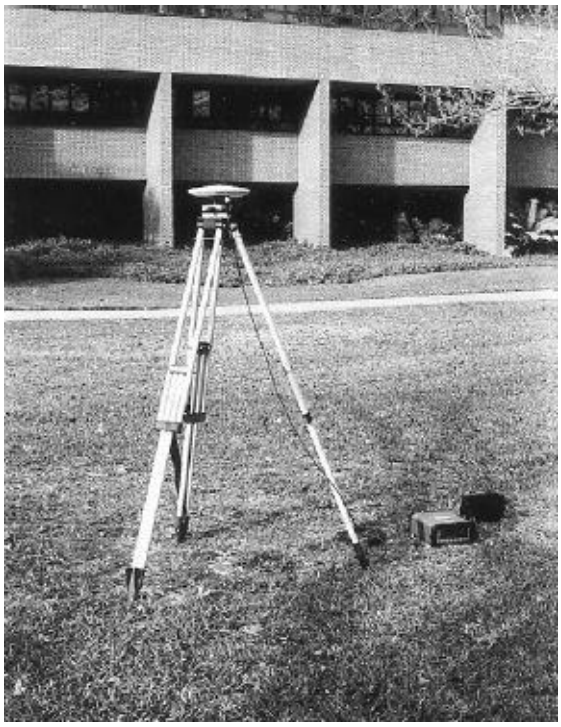


FIGURE 16-8 GPS receiver with antenna mounted on a tripod. (*Courtesy University of Florida.*)

The Department of Defense operates a series of ground tracking and control stations around the globe having GPS receivers which continuously monitor the entire constellation of satellites. Based on observations from these stations, accurate orbit information, given with respect to the WGS84 datum, is determined for each satellite. It is periodically transmitted to the satellites for subsequent broadcast as part of the GPS signal. This orbit information comprises what is called the *predicted ephemeris*. It enables GPS users to compute estimated positions of satellites through extrapolation from previously computed orbit parameters. A more accurate *precise ephemeris*, which contains directly computed parameters (as opposed to predicted), may be obtained from the National Geodetic Survey[1] several days after GPS observations have been obtained.

The fundamental mode of operation of GPS employs a single receiver and is called *point positioning*. In this mode, a receiver tracks the modulated code (C/A or P) from several satellites simultaneously. This information, the so-called code phase observable, is used to determine the time it takes for the signal to travel from the satellite to the receiver. The signal travel times from the various satellites are converted to distances known as *pseudoranges*, by multiplying by the speed of light. Given the positions of the satellites from the ephemeris, these distances are used in a three-dimensional spherical intersection to determine the coordinates of the receiver. Generally, at least four satellites are used which enable cancellation of the clock error of the receiver.

Errors inherent to point positioning are several. Most individually contribute an error of perhaps 1 to 5 m in the computed position of the receiver. These include errors due to ephemeris accuracy, timing accuracy, ionospheric and tropospheric interference, antenna construction, and multipath (or signal reflection). These errors are further amplified by a factor known as *PDOP* which is related to satellite geometry. The accumulation of all these errors can result in an error in receiver position of up to 25 m or more.

A GPS method known as *differential positioning* using code-phase observations can be used to determine locations of points with much greater accuracy than point positioning. The basic concept behind differential positioning is to employ two receivers which collect GPS signals simultaneously. One of the receivers is placed on a control point having known position, called a *base station;* and the other receiver, the *rover*, is placed at an unknown point. Since the position of the base station is known, ranging errors to individual satellites can be calculated. Since both receivers collected data simultaneously, the ranging errors calculated at the base station can be applied to the range measurements at the rover. Compensation for these ranging errors is made to compute more accurate coordinates of the unknown point.

For the ultimate in accuracy, *relative positioning* using carrier-phase observations is employed. In this approach, the phase changes of the L1 and unsually L2 carrier wave are measured at the receivers in order to determine their distances from the satellites. This is similar to the method employed by total station instruments in measuring distances electronically. The fundamental problem associated with this approach is that any particular cycle of the carrier wave is indistinguishable from the other cycles. The result is that there will be an unknown number of full cycles of the carrier wave between the satellite and the receiver. This *integer ambiguity* is a problem that must be overcome in the software used to process the information. Modern systems are able to overcome this problem very effectively. This is the preferred GPS method for establishing photogrammetric control.

Many modes of operation and data reduction techniques are available when performing ground control surveys by GPS. The method chosen will depend primarily upon accuracy requirements, but will also depend upon available equipment and processing software as well as the amount of time available to field crews. While an exhaustive list of GPS methods is not presented here, those that are commonly applied in photogrammetric control surveys are briefly discussed. The methods are broadly categorized into code-phase and carrier-phase techniques.

In general, single-receiver point positioning is far too inaccurate for useful photogrammetric control work. Therefore, differential techniques are used to establish control. If small-scale mapping is being performed, differential code-phase techniques may be employed. As mentioned previously, differential GPS requires that one receiver be placed at a known base station. One or more roving receivers may then be employed to collect simultaneous observations at unknown points. Depending on the quality of receivers used and the distance between the base station and rover, errors of less than 5 m can routinely be achieved. Currently the best-attainable accuracies of differential codephase GPS are better than 0.5 m.

Several carrier-phase methods of relative positioning are commonly used for establishing photogrammetric control. Of these, the *static* method is the most accurate. As the name implies, static GPS involves placing fixed receivers on points and collecting carrier-phase data for as long as an hour or more. After data have been collected, *baseline vectors* are computed between pairs of receivers which give ΔX, ΔY, and ΔZ components between corresponding points in a three-dimensional coordinate system. Generally, interconnected networks (see Fig. 16-1) of these vectors are created which are subsequently adjusted by least squares to obtain coordinates for the unknown points. Using the static technique, coordinates of unknown points can be determined with errors at the centimeter level.

If shorter observation times are desired, the method of *rapid static* observation may be employed. With this approach, station occupation times may be reduced to 10 min or even less, while still achieving centimeter-level accuracy. The rapid static method requires more expensive equipment and more sophisticated data processing techniques than static GPS, however. Generally, receivers capable of tracking both the L1 and L2 carrier frequencies as well as the C/A and P code signals are used. By

including all four phase measurements in a highly redundant solution, accurate coordinates can be determined despite the reduced station occupation time.

While GPS is most often used to compute horizontal position, it is capable of determining vertical position (elevation) to nearly the same level of accuracy. An inherent problem with the vertical position, however, is that it will be related to the ellipsoid, not the geoid or mean sea level (see Secs. 5-2 and 5-3). To relate the GPS-derived elevation (ellipsoid height) to the more conventional elevation (orthometric height), a geoid model is necessary (see Sec. 5-7).

16-6 Kinematic GPS Positioning

Another technique known as *kinematic positioning* can be used for establishing ground control. When performing kinematic positioning, the rover receiver is in constant motion except for the brief period during which a station is occupied. The method requires a stationary receiver to be located at a known base station, and an initial start-up period during which the rover receiver is held stationary. This start-up period is required so that the processing software can calculate the unknown integer ambiguities. Once the ambiguities have been resolved, the rover receiver can be moved, and as long as it remains in contact with the satellite transmissions, the changes in integer ambiguities can be determined. The accuracy of kinematic GPS can be nearly as good as that of static, i.e., within a few centimeters. A major drawback associated with kinematic GPS surveys is that loss of signal cannot be tolerated. When the rover receiver "loses lock" on the satellite transmissions, which can happen if it goes under a tree, bridge, or other obstruction, then it must be returned to a previously surveyed position so that the integer ambiguities can be redetermined. This restricts the use of kinematic GPS to open areas. To reduce the inconvenience associated with loss of lock, the technique of "on-the-fly" (OTF) ambiguity resolution has been developed. This method of ambiguity resolution is performed by a software technique whereby many trial combinations of integer ambiguities for the different satellites are tested in order to determine the correct set. This essentially trial-and-error approach generally requires a period of uninterrupted data (i.e., no loss of lock) of as much as several minutes in order for the ambiguities to be correctly determined. Use of dual-frequency receivers which track C/A as well as P code provides redundancy which can greatly enhance the computational process of determining the integer ambiguities.

Kinematic GPS can also be used for airborne GPS control. In this method, the position of the camera is measured using a GPS antenna fixed to the aircraft. Many of the same concepts from kinematic positioning for ground control apply to airborne control. For instance, a base station must be used and consequently multiple base stations must be used when mapping large areas to keep the distance to the station under a maximum threshold. Since loss of satellite lock is detrimental to the solution, care should be taken not to bank the aircraft to high roll angles during turns between flight lines in order to prevent the wings from blocking satellites. Typically, the maximum tolerated roll is about 15° to 20°. An important consideration regarding airborne GPS positioning is the synchronization (or lack thereof) of GPS fixes with the photographic exposures. The GPS receiver will be recording data at uniform time intervals called epochs, which may be on the order of 1 s each. The camera shutter, on the other hand, operates asynchronously with respect to the GPS fixes, as shown in Fig. 16-9. The result is that the position of the GPS antenna must first be interpolated from adjacent fixes before the coordinates can be translated to the lens. Depending upon atmospheric turbulence and the epoch recording rate of the GPS receiver, the error due to this interpolation can be quite severe. In Fig. 16-9, interpolation for exposures 1 and 2 results in sizable errors from the actual positions, whereas interpolation for exposure 4 is nearly perfect. Integration of airborne kinematic GPS and an inertial navigation system (INS) can greatly reduce the error from interpolating by "filling in the

gaps" between epochs. This method can provide position (and attitude) data at a rate of up to 200 Hz.

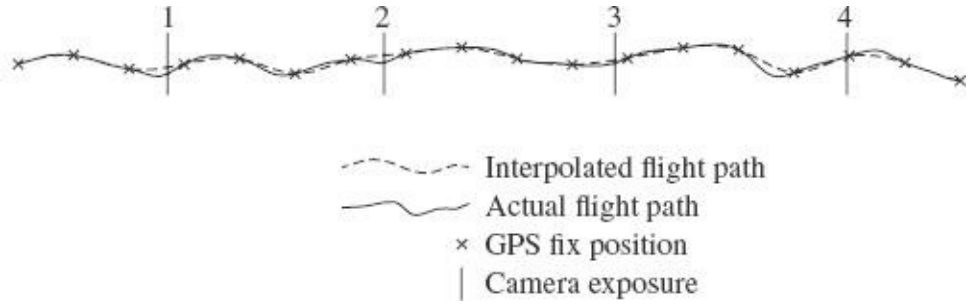


FIGURE 16-9 Interpolation between GPS fixes for positions at time of exposure.

16-7 Inertial Navigation Systems

For the most part, inertial navigation system (INS) technology has made possible the use of airborne laser scanning (LiDAR), linear array imaging sensors, and different types of radar and sonar for precise mapping. In the case of aerial photogrammetry, INS can provide a means for reducing or even eliminating the need for ground control. This is because, especially when integrated with airborne GPS, INS effectively enables the direct measurement of position and angular attitude associated with the exposure of a photograph. In other words, exterior orientation parameters, ω, φ, κ, X_L, Y_L, and Z_L, can be found without any information from the photography such as the image location of tie and/or control points. This is commonly referred to as *direct georeferencing*. The value of INS is clear when considering the potential reduction of the effort and cost of establishing ground control. INS is a complex subject and is given a simplified treatment in this section. For a more in-depth description, please see the references at the end of this chapter.

An INS consists of a computer and an inertial measurement unit (IMU). The IMU consists of a set of three *accelerometers* and three *gyroscopes* or *gyros*. The accelerometers are mounted orthogonally to each other and output directional *specific force*, which is the total acceleration of an object minus acceleration due to gravitational attraction. For example, an object stationary relative to the earth's surface has specific force of 9.8 m/s^2, whereas an object in freefall has zero specific force. Specific force is obtained from accelerometers via Newton's second law by measuring the amount of force required to keep an object stationary under different accelerations and dividing by the mass. There are multiple methods for mechanically implementing accelerometers. For instance, many modern devices use a coil in an electrical field to keep a mass in a stationary position. The gyros are also mounted orthogonally and are used to measure rotational motion. The main principle behind traditional gyros is conservation of angular momentum. One can imagine a toy gyroscope, a spinning top, or spinning plates. When no external force is applied to a rotating mass, its spin axis will remain constant. One can measure the torque required to keep a mounted gyro fixed with respect to a moving vehicle, and use it to calculate the vehicle's rate of angular change relative to the gyro's spin axis. Historically, rigid rotor gyros consisting of discs spinning around their center of mass were used for inertial-based navigation based on conservation of angular momentum. However, modern gyros, such as fiber optic and ring laser systems, are based on other physical laws. Although they operate on different principles, the fundamental output—the rate of angular change—is the same for modern gyros and spinning-mass gyros.

INS *mechanization* is a process for obtaining useable information from IMU measurements. Mechanization involves integrating the specific force from the accelerometers and the angular velocity from the gyros over time to calculate position, velocity, and attitude. An important

consideration in mechanization is the reference frame for the output parameters. IMU measurements are made with respect to the *inertial reference frame*. The origin of the inertial reference frame is the center of the earth. The *z* axis is parallel with the earth's spin axis. The *x* axis is oriented such that its positive end passes through the mean vernal equinox, the intersection of the plane of earth's orbit (the ecliptic) and the equator. The *y* axis completes the orthogonal system. Note that the geocentric coordinate system (see Sec. 5-4) rotates through the inertial reference frame once every 24 hours. The inertial measurements must be transformed into a mapping coordinate system such as the local-level frame, which is similar to the local vertical system in Sec. 5-5 except that the origin is located at the center of the INS. When mechanizing in the local vertical frame, the position is calculated in geodetic coordinates, while the velocity and attitude are with respect to easting, northing, and up. Since the mapping coordinate system is rotating with respect to the inertial reference frame, one must take into account not only gravity, but also the Coriolis effect and centripetal force. For a more in-depth description of INS mechanization, please see the references at the end of this chapter.

Since all output from the IMU is relative, one must first initialize the system before INS mechanization. This involves resolving an initial position, velocity, and attitude of the sensor. Although there are dynamic methods, it is simpler to initialize the INS when the platform is stationary. Initial position and velocity can be found using GPS. As for the attitude of the system, tilt can be determined by measuring the specific force when the system is not moving. The specific force measured will then only consist of gravity. The tilt of the system is the rotation that aligns the negative *z* axis of the sensor with the gravity vector (roughly down in the local-level frame). In order to initialize the heading one can use the gyro measurements when the vehicle is stationary, which are related to the rotation of the earth and the latitude of the initial position. The sensed rotation in the gyros corresponds to the transformed rotation about the *y* and *z* axes, east and north respectively, of the local-level frame relative to the inertial frame. More rigorous and precise methods usually follow the coarse method described here. In practice, the INS software performs initialization automatically with a specific protocol recommended by the manufacturer.

INS navigation where all positions are determined from a single initial position is called dead reckoning. Inertial sensors, while precise over short intervals, are susceptible to initial biases and random errors over time. Since in dead reckoning all subsequent measurements relate only to an initial state, errors accumulate in the course of collection. For this reason, the best measure of the quality of an IMU is the rate at which the gyros "drift" over time. The inherent accumulation of errors in position from INS mechanization has led to the widespread integration of GPS with INS.

16-8 GPS-INS Integration

Kinematic airborne GPS requires interpolation over sometimes significant lengths of time to estimate the position at exposure leading to unreliable values. In addition, GPS does not provide an estimate of angular attitude. INS can provide angular attitude, however, left unchecked INS estimates of position and attitude have unbounded errors over time. This has led to the development of integrated GPS-INS systems, and example application of which is shown in Fig. 16-10. In the figure, the black box mounted to the top of the sensor (in this case an airborne laser scanner) is the IMU. GPS and INS complement each other exceptionally well. INS provides high-frequency but volatile estimates, while GPS provides independent and relatively low-frequency estimates with bounded errors. Integrated GPS-INS systems use the GPS to update estimates from the INS. This could be achieved by simply using the GPS to "reset" the velocity and position from INS mechanization. However, integration is usually more rigorously accomplished using a *filter*—a mathematical process that provides best estimates of parameters given a function and input measurements. One such process is the *Kalman*

filter.

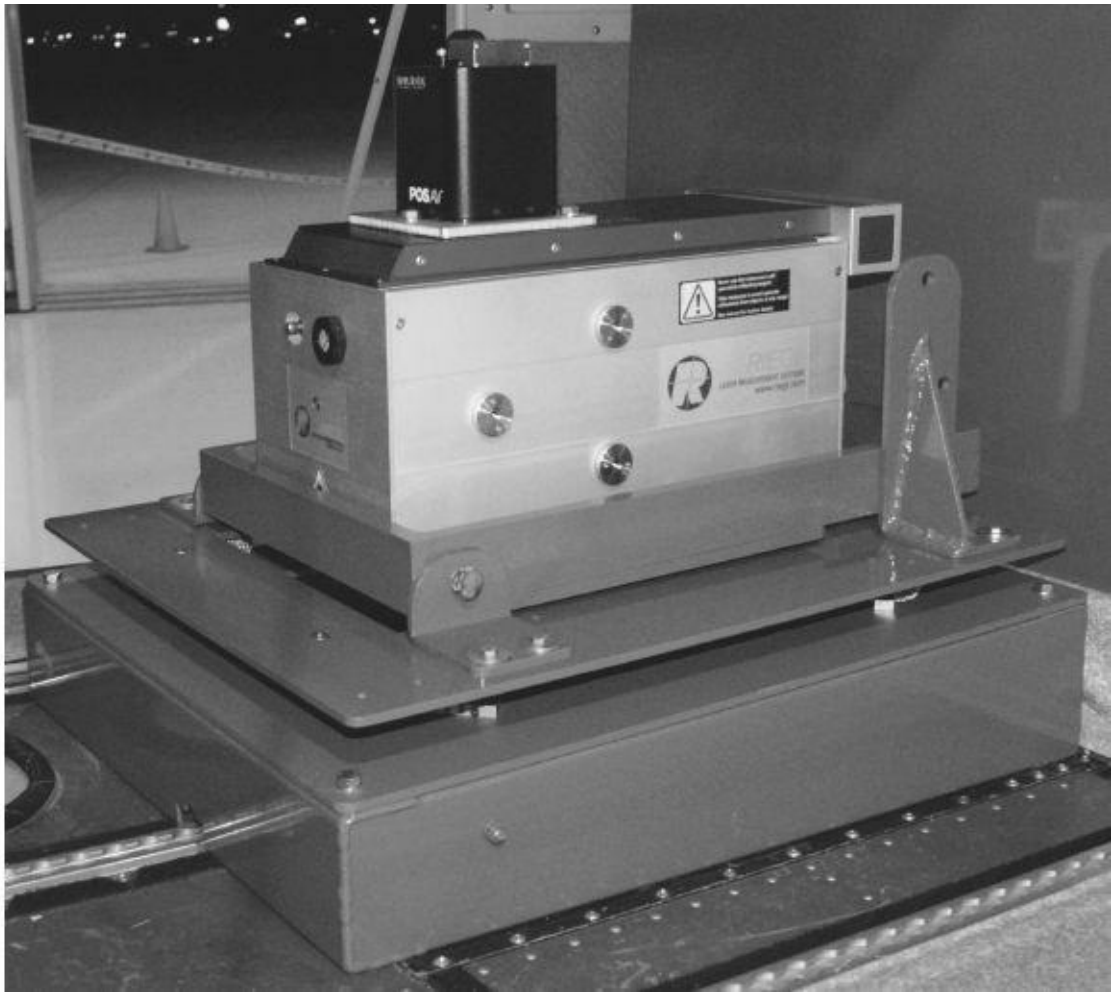


FIGURE 16-10 Airborne GPS-INS system. (*Courtesy of RIEGL USA, Inc.; http://www.rieglusa.com.*)

When integrating, the Kalman filter for INS-GPS operates on the errors associated with the INS. INS errors can be estimated through modeling of the system, and GPS enables a measure of the INS errors which can be used to refine their values. This "predict and update" scheme is characteristic of Kalman filters. It can be implemented at different levels of integration: *loosely coupled*, *tightly coupled*, and *deeply coupled*. Loosely coupled integration uses the independently calculated position and velocity from GPS. Whenever GPS position and velocity become available (at each epoch), they are subtracted from the values obtained from the INS. The difference between the two is used to calculate a new estimate of the error states which are then used to find new estimates of the position, velocity, and attitude. Tightly coupled integration incorporates more basic GPS measurements such as the pseudoranges to estimate the error states of the INS. Deeply coupled integration involves the incorporation of even more basic GPS observations, the satellite signals themselves, and requires special hardware to implement.

GPS-INS data is typically first processed using a filter operating forward in time. That is, each measurement is introduced into the filter sequentially as they were measured to produce the estimates. This means that each estimate is based on previous measurements, and only the very last estimate is derived from all the measurements. Following the forward filter, a backward-operating process is applied in order to allow forward-filtered estimates of position, velocity, and angular attitude to be adjusted based on subsequent estimates and measurements. This procedure is called *smoothing*.

In practice, integration of GPS-INS is a sophisticated process, and is an ongoing area of research

to develop new and more robust methods. It should be mentioned that other observations can be integrated with GPS and/or INS including those from altimeters, magnetometers, video, radar, and star-tracking devices among others. In fact, INS and vision-aided navigation in GPS-denied environments has become of particular interest.

Modern GPS-INS systems have become accurate and reliable. They can provide angular attitude accuracies as good as 0.005° rms for ω and φ, and 0.008° rms for κ. Positional accuracies for X_L, Y_L, and Z_L can be achieved at the level of 0.05 m rms. Furthermore, direct georeferencing of airborne imagery using integrated GPS-INS, in concert with elevation models provided by airborne laser scanning, allows for rapid production of photogrammetric products such as orthophotos. However, although direct georeferencing via GPS-INS is attractive due to its speed and efficiency, for best results it should be augmented with aerotriangulation. Aerotriangulation provides a check on the directly obtained exterior orientation parameters and can be implemented with directly georeferenced imagery with or without ground control points, although the former is preferred. The integration of aerotriangulation and airborne control is covered in Sec. 17-8. In addition, when using INS-GPS for airborne photogrammetry, it is necessary to know the relative attitude and position of the navigation sensors with respect to the camera. As shown in Sec. 17-8 aerotriangulation provides a method for calibrating these *boresight* and *lever arm* parameters.

References

Abidin, H. Z.: "On-the-Fly Ambiguity Resolution," *GPS World*, vol. 5, no. 4, 1994, p. 40.

American Society of Photogrammetry: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Anderson, J. M., and E. M. Mikhail: *Surveying: Theory and Practice*, WCB/McGraw-Hill, New York, 1998.

Barbour, N. and G. Schmidt.: "Inertial Sensor Technology Trends," *IEEE Sensors Journal*, vol. 1, no. 4, 2001, p. 332.

Bomford, G.: *Geodesy*, 4th ed., Clarendon Press, Oxford, 1980.

Ghilani, C. D., and P. R. Wolf: *Elementary Surveying: An Introduction to Geomatics*, 13th ed., Prentice Hall, New Jersey, 2011.

Ip, A., N. El-Sheimy, and M. Mostafa: "Performance Analysis of Integrated Sensor Orientation," *Photogrammetric Engineering and Remote Sensing*, vol. 73, no. 1, 2007, p. 1.

Leick, A.: *GPS Satellite Surveying*, 2d ed., Wiley, New York, 1995.

Mohamed, A. H. and A. Mamatas, in Thomas Lombaerts (ed.): *Flight Control*. Chapter: Fundamentals of GNSS-Aided Inertial Navigation, InTech—Open Access Publisher, ISBN 978-953-307-665-2, 2012.

Mohamed A. H. and K.P. Schwarz:"Adaptive Kalman Filtering for INS/GPS," *Journal of Geodesy*, vol. 73, no. 4, 1999, p. 193.

Schwarz, K. P., M. A. Chapman, M. E. Cannon, and P. Gong: "An Integrated IM/GPS Approach to the Georeferencing of Remotely Sensed Data," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 11, 1993, p. 1167.

Thompson, M. M.: *Maps for America*, U.S. Geological Survey, Government Printing Office, Washington, 1987.

VanWijk, M. C.: "Test Areas and Targeting in the Hull Project," *Canadian Surveyor*, vol. 25, no. 5, 1971, p. 514.

Wolf, P. R., and C. Ghilani: *Adjustment Computations: Statistics and Least Squares in Surveying and GIS*, Wiley, New York, 1997.

Problems

16-1. Explain the difference between basic control and photo control.

16-2. Describe the characteristics of good horizontal photo control points.

16-3. Describe the characteristics of good vertical photo control points.

16-4. State the national map accuracy standards for both horizontal positions and elevations.

16-5. Discuss the geospatial positioning accuracy standards, and contrast them with national map accuracy standards.

16-6. If a map is being prepared photogrammetrically to a scale of 600 ft/in, and photo control must be established to an accuracy 3 times greater than the allowable error for plotted points as specified by national map accuracy standards, how accurately on the ground must photo control points be located?

16-7. Repeat Prob. 16-6, except that map scale is 1 : 24,000.

16-8. What are the photo dimensions of the square at the intersection of two sidewalks of 2-m width if photo scale is 1 : 30,000?

16-9. What are the photographic dimensions in millimeters of a 0.75-m-diameter manhole cover if photo scale is 1 : 12,000?

16-10. Briefly describe three traditional field methods used in horizontal control surveys.

16-11. Give a brief description of two traditional field methods used in vertical control surveys.

16-12. Briefly describe the fundamental components of the Global Positioning System.

16-13. Explain the difference between carrier and code signals as it applies to GPS.

16-14. Discuss the errors which are inherent in single-receiver point positioning by GPS.

16-15. Briefly describe the principle of differential positioning by GPS. Why is this technique more accurate than point positioning?

16-16. Give a brief description of the concept of integer ambiguity as it applies to carrier-phase GPS.

16-17. Briefly discuss the principle of on-the-fly ambiguity resolution in kinematic GPS.

16-18. Discuss the advantages and disadvantages of using artificial targets as opposed to using natural targets for photo control.

16-19. What must be the ground dimension D (see Fig. 16-3) of artificial targets if their corresponding photo dimension is to be 0.05 mm on photos exposed from 3000 m above ground with a 152-mm-focal-length camera?

16-20. Repeat Prob. 16-19, except that the photo dimension is 10 pixels, flying height above ground is 6000 ft, and camera focal length is 13,000 measured in pixels.

16-21. Describe the measurements provided by inertial measurement units.

16-22. Explain the problems with "dead reckoning" with respect to an INS.

16-23. Describe the pros and cons of GPS and INS and why they are complementary systems.

16-24. Briefly describe the three levels of GPS-INS integration.

---

[1] Archival precise orbit information can be obtained at the following website: http://www.ngs.noaa.gov.

CHAPTER 17
Aerotriangulation

17-1 Introduction

Aerotriangulation is the term most frequently applied to the process of determining the *X*, *Y*, and *Z* ground coordinates of individual points based on photo coordinate measurements. *Phototriangulation* is perhaps a more general term, however, because the procedure can be applied to terrestrial photos as well as aerial photos. The principles involved are extensions of the material presented in Chap. 11. With improved photogrammetric equipment and techniques, accuracies to which ground coordinates can be determined by these procedures have become very high.

Aerotriangulation is used extensively for many purposes. One of the principal applications lies in extending or densifying ground control through strips and/or blocks of photos for use in subsequent photogrammetric operations. When used for this purpose, it is often called *bridging*, because in essence a "bridge" of intermediate control points is developed between field-surveyed control that exists in only a limited number of photos in a strip or block. Establishment of the needed control for compilation of topographic maps with stereoplotters is an excellent example to illustrate the value of aerotriangulation. In this application, as described in Chap. 12, the practical minimum number of control points necessary in each stereomodel is three horizontal and four vertical points. For large mapping projects, therefore, the number of control points needed is extensive, and the cost of establishing them can be extremely high if it is done exclusively by field survey methods. Much of this needed control is now routinely being established by aerotriangulation from only a sparse network of field-surveyed ground control and at a substantial cost savings. A more recent innovation involves the use of kinematic GPS and INS in the aircraft to provide coordinates and angular attitude of the camera at the instant each photograph is exposed. In theory, this method can eliminate the need for ground control entirely, although in practice a small amount of ground control is still used to strengthen the solution.

Besides having an economic advantage over field surveying, aerotriangulation has other benefits: (1) most of the work is done under laboratory conditions, thus minimizing delays and hardships due to adverse weather conditions; (2) access to much of the property within a project area is not required; (3) field surveying in difficult areas, such as marshes, extreme slopes, and hazardous rock formations, can be minimized; and (4) the accuracy of the field-surveyed control necessary for bridging is verified during the aerotriangulation process, and as a consequence, chances of finding erroneous control values after initiation of compilation are minimized and usually eliminated. This latter advantage is so meaningful that some organizations perform bridging even though adequate field-surveyed control exists for stereomodel control. It is for this reason also that some specifications for mapping projects require that aerotriangulation be used to establish photo control.

Apart from bridging for subsequent photogrammetric operations, aerotriangulation can be used in a variety of other applications in which precise ground coordinates are needed although most of these uses have been largely supplanted by GPS. In property surveying, aerotriangulation can be used to locate section corners and property corners or to locate evidence that will assist in finding these corners. In topographic mapping, aerotriangulation can be used to develop digital elevation models by computing *X*, *Y*, and *Z* ground coordinates of a systematic network of points in an area, although airborne laser scanning is commonly being used for this task. Aerotriangulation has been used

successfully for densifying geodetic control networks in areas surrounded by tall buildings where problems due to multipath cause a loss of accuracy in GPS surveys. Special applications include the precise determination of the relative positions of large machine parts during fabrication. It had been found especially useful in such industries as shipbuilding and aircraft manufacture. Many other applications of aerotriangulation are also being pursued.

Methods of performing aerotriangulation may be classified into one of three categories: *analog*, *semianalytical*, and *analytical*. Early analog procedures involved manual interior, relative, and absolute orientation of the successive models of long strips of photos using stereoscopic plotting instruments having several projectors. This created long strip models from which coordinates of pass points could be read directly. Later, universal stereoplotting instruments were developed which enabled this process to be accomplished with only two projectors. These procedures are now principally of historical interest, having given way to the other two methods.

Semianalytical aerotriangulation involves manual interior and relative orientation of stereomodels within a stereoplotter, followed by measurement of model coordinates. Absolute orientation is performed numerically—hence the term *semianalytical* aerotriangulation.

Analytical methods consist of photo coordinate measurement followed by numerical interior, relative, and absolute orientation from which ground coordinates are computed. Various specialized techniques have been developed within each of the three aerotriangulation categories. This chapter briefly describes some of these techniques. It predominantly relates to bridging for subsequent photogrammetric operations because this is the principal use of aerotriangulation. Extension of these basic principles can readily be translated to the other areas of application, however.

17-2 Pass Points for Aerotriangulation

Pass points for aerotriangulation are normally selected in the general photographic locations shown in Fig. 17-1*a*. Historically, points were artificially generated using stereoscopic point marking devices. These devices involved drilling a hole in the photograph, destroying the emulsion on that point. Nowadays, in automatic aerotriangulation, pass points are usually found using automated procedures on digital and scanned-film photography. Control points can be located manually with sub-pixel accuracy in software. There is no destruction of the photograph using digital methods, so points can easily be removed and replaced.

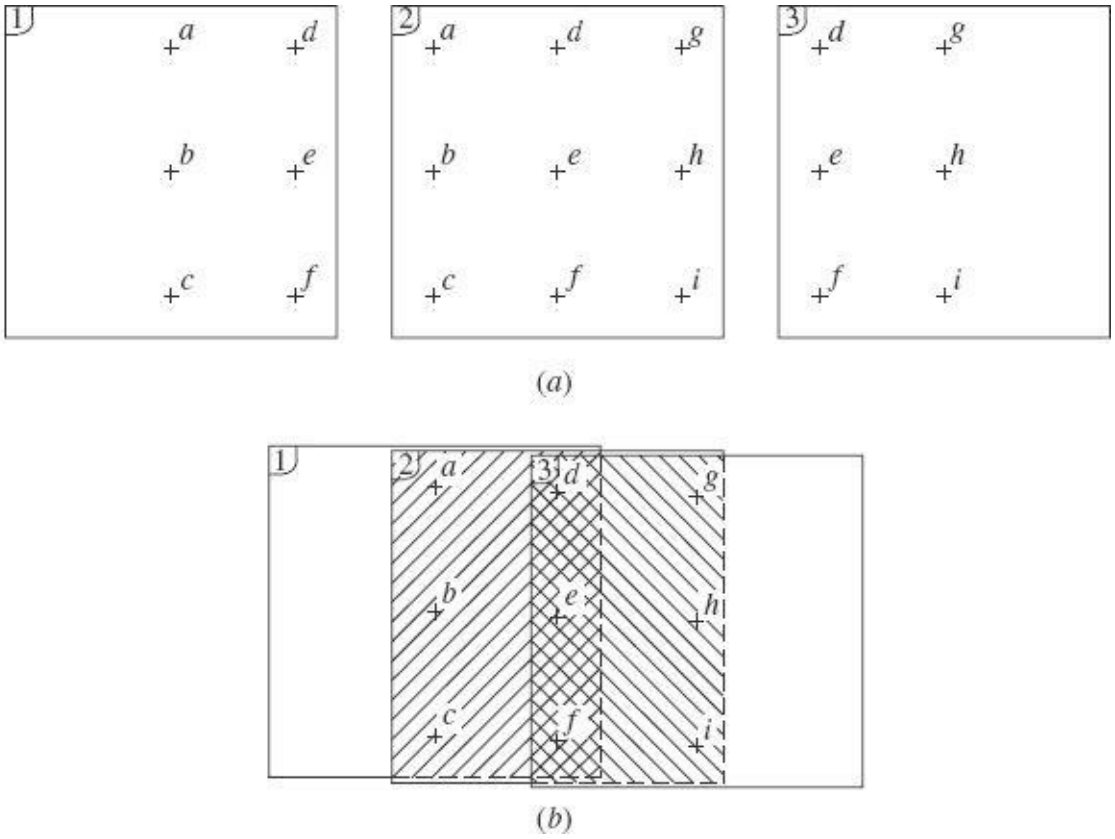


FIGURE 17-1 (*a*) Idealized pass point locations for aerotriangulation. (*b*) Locations of pass points in two adjacent stereomodels.

A typical procedure for measuring a pass point begins by first manually digitizing the point in one photograph. The pixels around this point serve as the template array. Next, the user defines a search area in other photographs for automatic image matching. There are also automatic methods for defining a search area by predicting the coordinates of the point in the subsequent photographs. Finally, the pixel patch in the search area corresponding to the template array is automatically located. Normalized cross-correlation followed by least squares matching is a common method for this step (see Sec. 15-8). To avoid poor matches and blunders, well-defined unique objects with good contrast and directionality should be selected as image-matching templates. Image-matching software usually provides a measure of how well the point was matched, such as the correlation coefficient in normalized cross-correlation. This number should serve as a guide for the user to decide whether or not to accept the matching results. Care must be taken because it is not uncommon for incorrectly matched points to have high correlation coefficients. The process is repeated for each pass point keeping in mind the optimal distribution illustrated in Fig. 17-1. Due to increased redundancy, the most effective points are those that appear in the so-called *tri-lap* area, which is the area included on three consecutive images along a strip. Once many pass points are located, more can be added in a fully automated process by prediction of point locations based on a coordinate transformation.

17-3 Fundamentals of Semianalytical Aerotriangulation

Semianalytical aerotriangulation, often referred to as *independent model* aerotriangulation, is a partly analytical procedure that emerged with the development of computers. It involves relative orientation of each stereomodel of a strip or block of photos. After the models have been formed, they are numerically adjusted to the ground system by either a sequential or a simultaneous method. In the sequential approach, contiguous models are joined analytically, one by one, to form a continuous strip

model, and then absolute orientation is performed numerically to adjust the strip model to ground control. In the simultaneous approach, all models in a strip or block are joined and adjusted to ground control in a single step, much like the simultaneous transformation technique described in Sec. 9-8.

An advantage of using semianalytical aerotriangulation is that independent stereomodels are more convenient for operators in production processes. This stems from the fact that the images that make up stereomodel are more "tightly" oriented with respect to each other, whereas in fully analytical adjustments the images are oriented to optimize their fit with respect to a block of multiple photos which may lead to residual y parallax in the orientation between individual stereopairs. Regardless of whether the sequential or simultaneous method is employed, the process yields coordinates of the pass points in the ground system. Additionally, coordinates of the exposure stations can be determined in either process; Thus, semianalytical solutions can provide initial approximations for a subsequent bundle adjustment (see Sec. 17-6).

17-4 Sequential Construction of a Strip Model from Independent Models

In the sequential approach to semianalytical aerotriangulation, each stereopair of a strip is relatively oriented in a stereoplotter, the coordinate system of each model being independent of the others. When relative orientation is completed, model coordinates of all control points and pass points are read and recorded. This is done for each stereomodel in the strip. Figures 17-2*a* and *b* illustrate the first three relatively oriented stereomodels of a strip and show plan views of their respective independent coordinate systems. By means of pass points common to adjacent models, a three-dimensional conformal coordinate transformation (see Sec. C-7) is used to tie each successive model to the previous one. To gain needed geometric strength in the transformations, the coordinates of the perspective centers (model exposure stations) are also measured in each independent model and included as common points in the transformation. The right exposure station of model 1-2, O_2, for example, is the same point as the left exposure station of model 2-3. To transform model 2-3 to model 1-2, therefore, coordinates of common points d, e, f, and O_2 of model 2-3 are made to coincide with their corresponding model 1-2 coordinates. Once the parameters for this transformation have been computed, they are applied to the coordinates of points g, h, i, and O_3 in the system of model 2-3 to obtain their coordinates in the model 1-2 system. These points in turn become control for a transformation of the points of model 3-4. By applying successive coordinate transformations, a continuous strip of stereomodels may be formed, as illustrated in Fig. 17-2*c*. The entire strip model so constructed is in the coordinate system defined by model 1-2.

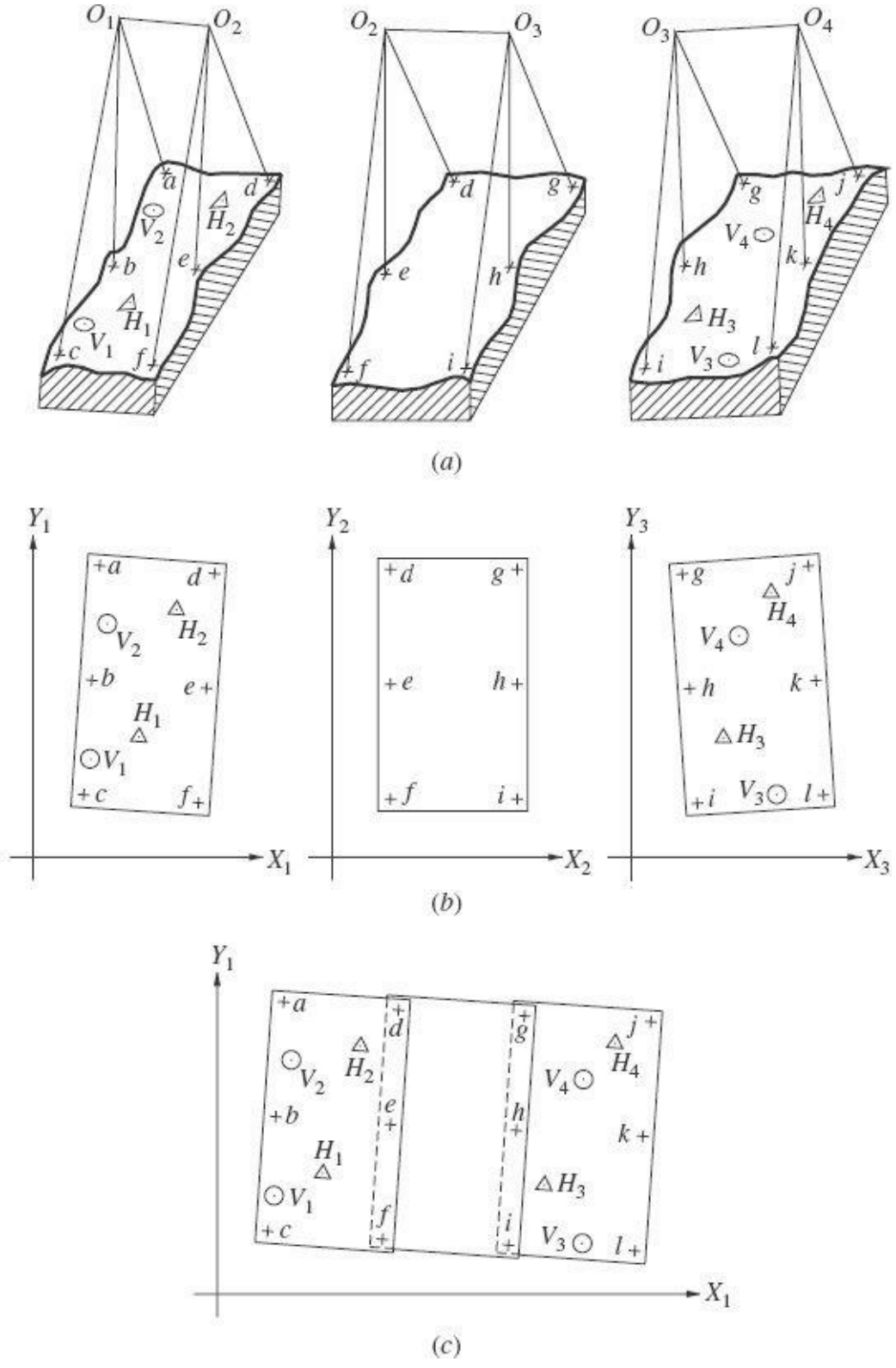


FIGURE 17-2 Independent model or semianalytical aerotriangulation. (*a*) Three adjacent relatively oriented stereomodels. (*b*) Individual arbitrary coordinate systems of three adjacent stereomodels. (*c*) Continuous strip of stereomodels formed by numerically joining the individual arbitrary coordinate systems into one system.

17-5 Adjustment of a Strip Model to Ground

After a strip model has been formed, it is numerically adjusted to the ground coordinate system using all available control points. If the strip is short, i.e., up to about four models, this adjustment may be done using a three-dimensional conformal coordinate transformation. This requires that a minimum of

two horizontal control points and three vertical control points be present in the strip. More control than the minimum is desirable, however, as it adds stability and redundancy to the solution. As discussed later in this section, if the strip is long, a polynomial adjustment is preferred to transform model coordinates to the ground coordinate system. In the short strip illustrated in Fig. 17-2*c*, horizontal control points H_1 through H_4 and vertical control points V_1 through V_4 would be used in a three-dimensional conformal coordinate transformation to compute the ground coordinates of pass points *a* through *l* and exposure stations O_1 through O_4.

Example 17-1

Figure 17-3 illustrates a continuous strip of three stereomodels with pass points *a* through *l* and ground control points *A* through *E*. Independent model coordinates for points and exposure stations of each model are listed below along with ground coordinates of control points *A* through *E*. Compute the ground coordinates of the pass points and exposure stations by the sequential method of semianalytical aerotriangulation. Use a three-dimensional conformal coordinate transformation program.

Model Coordinates (mm)											
Model 1-2				Model 2-3				Model 3-4			
Point	x	y	z	Point	x	y	z	Point	x	y	z
a	−1.568	91.538	−0.243	d	−0.120	82.885	9.473	g	−0.869	104.502	−12.763
b	−3.995	−0.725	−1.124	e	−6.933	−5.499	10.627	h	0.623	9.411	−16.914
c	−2.453	−88.126	−3.378	f	−5.356	−84.482	9.860	i	3.259	−86.164	−18.112
d	92.922	92.610	−0.588	g	78.090	83.401	8.690	j	98.487	98.233	−12.522
e	81.781	−1.480	−1.881	h	80.939	2.593	7.601	k	100.807	3.273	−15.224
f	80.019	−85.903	−4.881	i	84.843	−78.510	8.957	l	98.299	−98.052	−16.039
A	23.295	65.291	−0.620	C	44.480	20.499	8.869	D	62.943	101.805	−12.672
B	34.886	−65.717	−3.919	O_2	0.000	0.000	152.292	E	83.271	−56.709	−15.386
O_1	0.000	0.000	152.292	O_3	85.079	−0.955	151.424	O_4	0.000	0.000	152.292

Ground Control Coordinates (m)			
Point	X	Y	Z
A	1444.632	2892.290	88.904
B	1525.490	2109.281	83.661
C	2133.597	2655.634	89.583
D	2705.480	3035.452	92.118
E	2811.033	2173.558	94.052

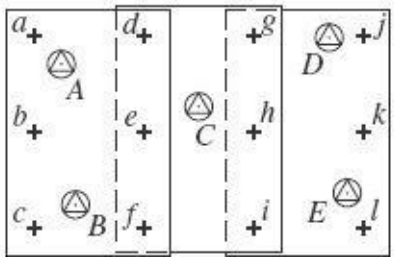


FIGURE 17-3 Configuration of pass points and control for semianalytical aerotriangulation of

Example 17-1.

Solution

1. With an ASCII text editor, create the following data file with a ".dat" extension (see Example 11-3 for a description of the data file format):

```
d    -0.120   82.885    9.473  92.922   92.610   -0.588
e    -6.933   -5.499   10.627  81.781   -1.480   -1.881
f    -5.356  -84.482    9.860  80.019  -85.903   -4.881
O2    0.000    0.000  152.292  89.009    0.223  149.734
#
g    78.090   83.401    8.690
h    80.939    2.593    7.601
i    84.843  -78.510    8.957
C    44.480   20.499    8.869
O3   85.079   -0.955  151.424
#
```

2. Run the "3dconf" program to produce the following results. (Only the portion of the output which gives transformed points is shown.)

```
Transformed Points:
 Point      X         Y         Z      SDev.X  SDev.Y  SDev.Z
     g   176.501    89.791   -1.310   0.037   0.036   0.035
     h   176.019     3.397   -4.702   0.027   0.032   0.031
     i   176.645   -83.426   -5.488   0.036   0.036   0.036
     C   137.848    24.079   -2.920   0.024   0.025   0.024
    O3   179.857    -4.535  148.932   0.040   0.038   0.038
```

The above output gives the coordinates of points *g*, *h*, *i*, *C*, and O_3 in the model 1-2 system.

3. With an ASCII text editor, create the following data file with a ".dat" extension:

```
g    -0.869  104.502  -12.763  176.501   89.791   -1.310
h     0.623    9.411  -16.914  176.019    3.397   -4.702
i     3.259  -86.164  -18.112  176.645  -83.426   -5.488
O3    0.000    0.000  152.292  179.857   -4.535  148.932
#
j    98.487   98.233  -12.522
k   100.807    3.273  -15.224
l    98.299  -98.052  -16.039
D    62.943  101.805  -12.672
E    83.271  -56.709  -15.386
O4   96.068   -1.067  152.499
#
```

4. Run the "3dconf" program to produce the following results. (Only the portion of the output which gives transformed points is shown.)

```
Transformed Points:
 Point      X         Y         Z      SDev.X  SDev.Y  SDev.Z
     j   266.567    82.280   -3.733   0.010   0.010   0.010
     k   266.877    -3.993   -5.860   0.008   0.009   0.009
     l   262.742   -95.950   -6.117   0.011   0.010   0.010
     D   234.369    86.171   -2.926   0.009   0.009   0.008
     E   249.870   -58.135   -5.288   0.009   0.009   0.008
    O4   267.032    -7.252  146.536   0.011   0.011   0.010
```

The output gives the coordinates of points *j*, *k*, *l*, *D*, *E*, and O_4 in the model 1-2 system.

5. With an ASCII text editor, create the following data file with a ".dat" extension:

```
A    23.295    65.291    -0.620   1444.632   2892.290   88.904
B    34.886   -65.717    -3.919   1525.490   2109.281   83.661
C   137.848    24.079    -2.920   2133.597   2655.634   89.583
D   234.369    86.171    -2.926   2705.480   3035.452   92.118
E   249.870   -58.135    -5.288   2811.033   2173.558   94.052
#
a    -1.568    91.538    -0.243
b    -3.995    -0.725    -1.124
c    -2.453   -88.126    -3.378
d    92.922    92.610    -0.588

e    81.781    -1.480    -1.881
f    80.019   -85.903    -4.881
O1    0.000     0.000   152.292
O2   89.009     0.223   149.734
g   176.501    89.791    -1.310
h   176.019     3.397    -4.702
i   176.645   -83.426    -5.488
O3  179.857    -4.535   148.932
j   266.567    82.280    -3.733
k   266.877    -3.993    -5.860
l   262.742   -95.950    -6.117
O4  267.032    -7.252   146.536
#
```

6. Run the "3dconf" program to produce the following results (only the portion of the output which gives transformed points is shown):

Transformed Points:

Point	X	Y	Z	SDev.X	SDev.Y	SDev.Z
a	1293.555	3047.090	86.231	0.062	0.062	0.080
b	1287.111	2494.926	90.169	0.057	0.057	0.064
c	1304.113	1972.055	85.755	0.065	0.065	0.084
d	1858.691	3061.799	92.594	0.047	0.047	0.062
e	1800.320	2497.897	93.466	0.040	0.040	0.042
f	1797.369	1992.472	83.989	0.050	0.050	0.069
O1	1296.858	2515.085	1008.088	0.081	0.106	0.080
O2	1829.486	2524.009	1000.806	0.070	0.097	0.063
g	2358.935	3052.234	96.113	0.046	0.046	0.060
h	2363.847	2535.093	84.608	0.038	0.038	0.039
i	2375.185	2015.747	88.834	0.048	0.048	0.067
O3	2373.385	2503.479	1004.702	0.070	0.096	0.063
j	2898.543	3015.005	90.524	0.059	0.059	0.073
k	2908.066	2498.785	86.644	0.055	0.055	0.062
l	2891.322	1948.364	94.128	0.064	0.064	0.087
O4	2895.283	2494.672	998.523	0.079	0.103	0.078

This completes the solution. Note that the output of step 6 contains the computed ground coordinates in meters for pass points *a* through *l* as well as the exposure stations O_1 through O_4. The computed standard deviations are also listed in meters. ▲

Due to the nature of sequential strip formation, random errors will accumulate along the strip. Often, this accumulated error will manifest itself in a systematic manner with the errors increasing in a nonlinear fashion. This effect, illustrated in Fig. 17-4, can be significant, particularly in long strips. Figure 17-4*a* shows a strip model comprised of seven contiguous stereomodels from a single flight line. Note from the figure that sufficient ground control exists in model 1-2 to absolutely orient it (and thereby the entire strip) to the ground system. The remaining control points (in models 4-5 and 7-8)

can then be used as *checkpoints* to reveal accumulated errors along the strip. Figure 17-4*b* shows a plot of the discrepancies between model and ground coordinates for the checkpoints as a function of X coordinates along the strip. Except for the ground control in the first model, which was used to absolutely orient the strip, discrepancies exist between model positions of horizontal and vertical control points and their corresponding field-surveyed positions. Smooth curves are fit to the discrepancies as shown in the figure.

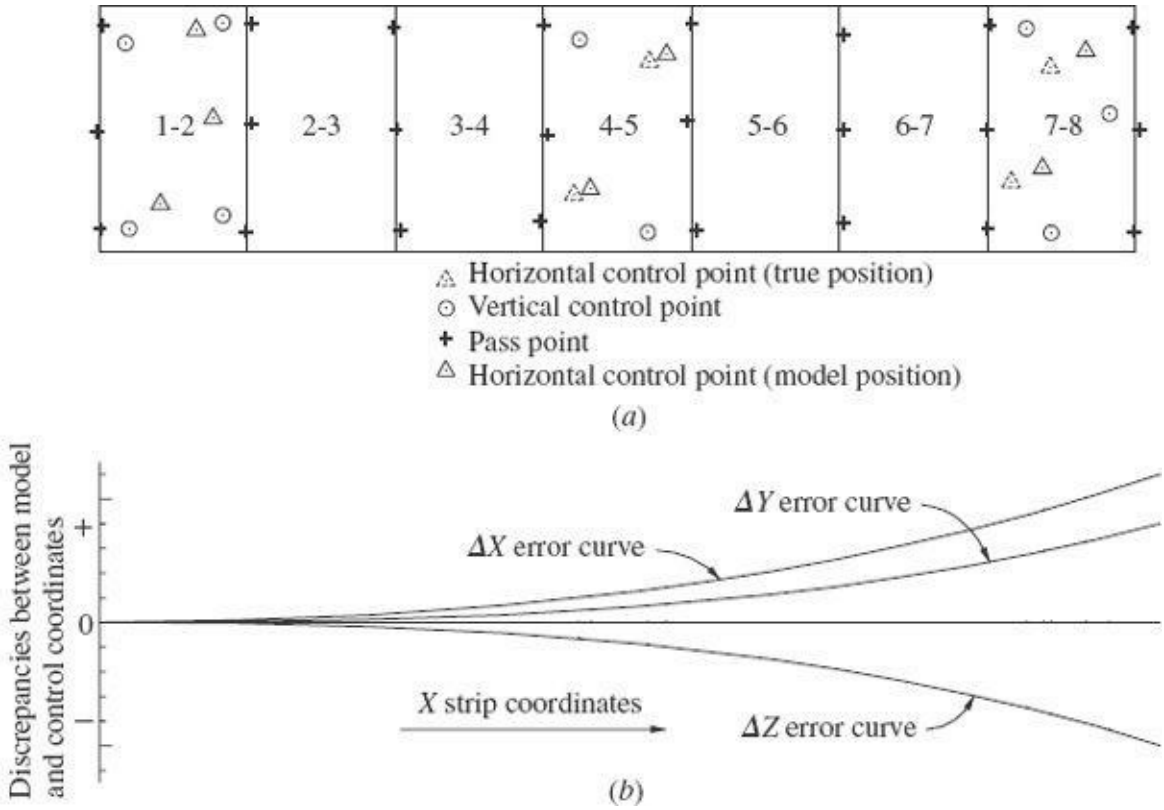


FIGURE 17-4 (*a*) Plan view of control extension of a seven-model strip. (*b*) Smooth curves indicating accumulation of errors in X, Y, and Z coordinates during control extension of a strip.

If sufficient control is distributed along the length of the strip, a three-dimensional polynomial transformation can be used in lieu of a conformal transformation to perform absolute orientation and thus obtain corrected coordinates for all pass points. This polynomial transformation yields higher accuracy through modeling of systematic errors along the strip. Most of the polynomials in use for adjusting strips formed by aerotriangulation are variations of the following third-order equations:

$$\begin{aligned}
\overline{X} &= a_0 + a_1X + a_2Y + a_3X^2 + a_4XY + a_5Y^2 + a_6X^3 + a_7X^2Y + a_8XY^2 + a_9Y^3 \\
\overline{Y} &= b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + b_6X^3 + b_7X^2Y + b_8XY^2 + b_9Y^3 \\
\overline{Z} &= c_0 + c_1X + c_2Y + c_3X^2 + c_4XY + c_5Y^2 + c_6X^3 + c_7X^2Y + c_8XY^2 + c_9Y^3
\end{aligned} \tag{17-1}$$

In Eqs. (17-1), $\overline{X}$, $\overline{Y}$, and $\overline{Z}$ are the transformed ground coordinates; X and Y are strip model coordinates; and the a's, b's, and c's are coefficients which define the shape of the polynomial error curves. The equations contain 30 unknown coefficients (a's, b's, and c's). Each three-dimensional control point enables the above three polynomial equations to be written, and thus 10 three-dimensional control points are required in the strip for an exact solution. When dealing with

transformations involving polynomials, however, it is imperative to use redundant control which is well distributed throughout the strip. It is important that the control points occur at the periphery as well, since extrapolation from polynomials can result in excessive corrections. As illustrated by Fig. 17-4*b*, errors in *X*, *Y*, and *Z* are principally functions of the linear distance (*X* coordinate) of the point along the strip. However, the nature of error propagation along strips formed by aerotriangulation is such that discrepancies in *X*, *Y*, and *Z* coordinates are also each somewhat related to the *Y* positions of the points in the strip. Depending on the complexity of the distortion, certain terms may be eliminated from Eqs. (17-1) if they are found not to be significant. This serves to increase redundancy in the transformation which generally results in more accurate results. Further, the discussion of the application of polynomials in adjusting strip models to ground can be found in references cited at the end of this chapter. It is possible however, to avoid the polynomial adjustment completely by using the simultaneous approach as mentioned in Sec. 17-3.

17-6 Simultaneous Bundle Adjustment

The most elementary approaches to analytical aerotriangulation consist of the same basic steps as those of analog and semianalytical methods and include (1) relative orientation of each stereomodel, (2) connection of adjacent models to form continuous strips and/or blocks, and (3) simultaneous adjustment of the photos from the strips and/or blocks to field-surveyed ground control. What is different about analytical methods is that the basic input consists of precisely measured photo coordinates of control points and pass points. Relative orientation is then performed analytically based upon the measured coordinates and known camera constants. Finally, the entire block of photographs is adjusted simultaneously to the ground coordinate system.

Analytical aerotriangulation tends to be more accurate than analog or semianalytical methods, largely because analytical techniques can more effectively eliminate systematic errors such as film shrinkage, atmospheric refraction distortions, and camera lens distortions. In fact, *X* and *Y* coordinates of pass points can quite routinely be located analytically to an accuracy of within about 1/15,000 of the flying height, and *Z* coordinates can be located to an accuracy of about 1/10,000 of the flying height. With specialized equipment and procedures, planimetric accuracy of 1/350,000 of the flying height and vertical accuracy of 1/180,000 have been achieved. Another advantage of analytical methods is the freedom from the mechanical or optical limitations imposed by stereoplotters. Photography of any focal length, tilt, and flying height can be handled with the same efficiency. The calculations involved are rather complex; however, a number of suitable computer programs are available to perform analytical aerotriangulation. With the advances in computer power, these calculations have become routine.

Several different variations in analytical aerotriangulation techniques have evolved. Basically, however, all methods consist of writing condition equations that express the unknown elements of exterior orientation of each photo in terms of camera constants, measured photo coordinates, and ground coordinates. The equations are solved to determine the unknown orientation parameters, and either simultaneously or subsequently, coordinates of pass points are calculated. By far the most common condition equations used are the collinearity equations which are presented in Sec. 11-4 and derived in App. D. Analytical procedures have been developed which can simultaneously enforce collinearity conditions onto units which consist of hundreds of photographs.

The ultimate extension of the principles described in the preceding sections is to adjust all photogrammetric measurements to ground control values in a single solution known as a *bundle adjustment*. The process is so named because of the many light rays that pass through each lens position constituting a bundle of rays. The bundles from all photos are adjusted simultaneously so that

corresponding light rays intersect at positions of the pass points and control points on the ground. The process is an extension of the principles of analytical photogrammetry presented in Chap. 11, applied to an unlimited number of overlapping photographs.

Figure 17-5*a* shows a small block consisting of two strips with four photos per strip. The photo block contains images of 20 pass points labeled 1 through 20 and 6 control points labeled *A* through *F*, for a total of 26 object points. Points 3, 8, 13, 18, *B*, and *E* also serve as *tie points* which connect the two adjacent strips. Figure 17-5*b* shows the individual photos in a nonoverlapping configuration. Note that photos 1, 4, 5, and 8 each contain images of 8 points; and photos 2, 3, 6, and 7 each contain images of 11 points, for a grand total of $4 \times 8 + 4 \times 11 = 76$ point images.

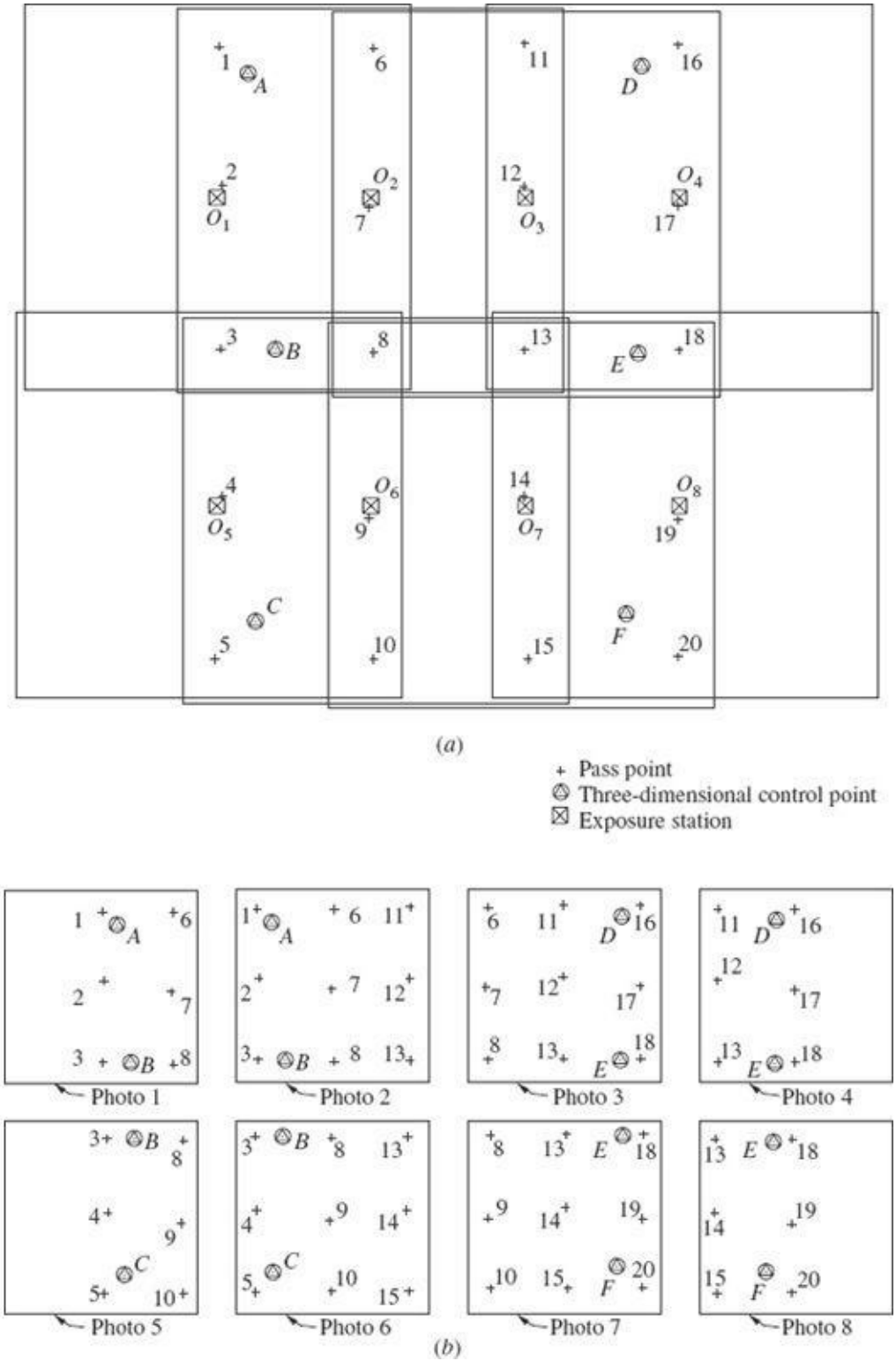


FIGURE 17-5 (*a*) Block of photos in overlapped position. (*b*) Separated photos showing image points.

The unknown quantities to be obtained in a bundle adjustment consist of (1) the X, Y, and Z object space coordinates of all object points and (2) the exterior orientation parameters (ω, ϕ, κ, X_L, Y_L, and Z_L) of all photographs. The first group of unknowns (object space coordinates) is the necessary result of any aerotriangulation, analytical or otherwise. Exterior orientation parameters, however, are generally not of interest to the photogrammetrist, but they must be included in the mathematical model for consistency. In the photo block of Fig. 17-5*a* the number of unknown object coordinates is 26 × 3 = 78 (number of object points times the number of coordinates per point). The number of unknown exterior orientation parameters is 8 × 6 = 48 (number of photos times the number of exterior orientation parameters per photo). Therefore the total number of unknowns is 78 + 48 = 126.

The measurements (observed quantities) associated with a bundle adjustment are (1) x and y photo coordinates of images of object points; (2) X, Y, and/or Z coordinates of ground control points; and (3) direct observations of the exterior orientation parameters (ω, ϕ, κ, X_L, Y_L, and Z_L) of the photographs. The first group of observations, photo coordinates, is the fundamental photogrammetric measurements. For a proper bundle adjustment they need to be weighted according to the accuracy and precision with which they were measured. The next group of observations is coordinates of control points determined through field survey. Although ground control coordinates are indirectly determined quantities, they can be included as observations provided that proper weights are assigned. The final set of observations, exterior orientation parameters, has recently become important in bundle adjustments with the use of airborne GPS control as well as *inertial navigation systems* (INSs) which have the capability of measuring the angular attitude of a photograph.

Returning to the block of Fig. 17-5, the number of photo coordinate observations is 76 × 2 = 152 (number of imaged points times the number of photo coordinates per point), and the number of ground control observations is 6 × 3 = 18 (number of three-dimensional control points times the number of coordinates per point). If the exterior orientation parameters were measured, the number of additional observations would be 8 × 6 = 48 (number of photos times the number of exterior orientation parameters per photo). Thus, if all three types of observations are included, there will be a total of 152 + 18 + 48 = 218 observations; but if only the first two types are included, there will be only 152 + 18 = 170 observations. Regardless of whether exterior orientation parameters were observed, a least squares solution is possible since the number of observations is greater than the number of unknowns (126) in either case.

The observation equations which are the foundation of a bundle adjustment are the collinearity equations (see Sec. D-3). These equations are given below in a slightly modified form as Eqs. (17-2) and (17-3).

$$x_{ij} = x_o - f\left[\frac{m_{11_i}(X_j - X_{L_i}) + m_{12_i}(Y_j - Y_{L_i}) + m_{13_i}(Z_j - Z_{L_i})}{m_{31_i}(X_j - X_{L_i}) + m_{32_i}(Y_j - Y_{L_i}) + m_{33_i}(Z_j - Z_{L_i})}\right] \tag{17-2}$$

$$y_{ij} = y_o - f\left[\frac{m_{21_i}(X_j - X_{L_i}) + m_{22_i}(Y_j - Y_{L_i}) + m_{23_i}(Z_j - Z_{L_i})}{m_{31_i}(X_j - X_{L_i}) + m_{32_i}(Y_j - Y_{L_i}) + m_{33_i}(Z_j - Z_{L_i})}\right] \tag{17-3}$$

In these equations, x_{ij} and y_{ij} are the measured photo coordinates of the image of point j on photo i related to the fiducial axis system; x_o and y_o are the coordinates of the principal point in the fiducial axis system; f is the focal length (or more correctly, principal distance) of the camera;

$m_{11_i}, m_{12_i}, \ldots, m_{33_i}$ are the rotation matrix terms for photo i; X_j, Y_j, and Z_j are the coordinates of point j in object space; and X_{L_i}, Y_{L_i}, and Z_{L_i} are the coordinates of the incident nodal point of the camera lens in object space. Since the collinearity equations are nonlinear, they are linearized by applying the first-order terms of Taylor's series at a set of initial approximations. After linearization (see Sec. D-5) the equations can be expressed in the following matrix form:

$$\dot{B}_{ij}\dot{\Delta}_i + \ddot{B}_{ij}\ddot{\Delta}_j = \varepsilon_{ij} + V_{ij} \tag{17-4}$$

where

$$\dot{B}_{ij} = \begin{bmatrix} b_{11_{ij}} & b_{12_{ij}} & b_{13_{ij}} & -b_{14_{ij}} & -b_{15_{ij}} & -b_{16_{ij}} \\ b_{21_{ij}} & b_{22_{ij}} & b_{23_{ij}} & -b_{24_{ij}} & -b_{25_{ij}} & -b_{26_{ij}} \end{bmatrix} \quad \ddot{B}_{ij} = \begin{bmatrix} b_{14_{ij}} & b_{15_{ij}} & b_{16_{ij}} \\ b_{24_{ij}} & b_{25_{ij}} & b_{26_{ij}} \end{bmatrix}$$

$$\dot{\Delta}_i = \begin{bmatrix} d\omega_i \\ d\phi_i \\ d\kappa_i \\ dX_{L_i} \\ dY_{L_i} \\ dZ_{L_i} \end{bmatrix} \quad \ddot{\Delta}_j = \begin{bmatrix} dX_j \\ dY_j \\ dZ_j \end{bmatrix} \quad \varepsilon_{ij} = \begin{bmatrix} J_{ij} \\ K_{ij} \end{bmatrix} \quad V_{ij} = \begin{bmatrix} v_{x_{ij}} \\ v_{y_{ij}} \end{bmatrix}$$

The above terms are defined as for Eqs. (D-15) and (D-16), except that subscripts i and j are used for the photo designation and point designation, respectively. (In App. D, the subscript A is used for the point designation, and no subscript is used for the photo designation.) Matrix $\dot{B}_{ij}$ contains the partial derivatives of the collinearity equations with respect to the exterior orientation parameters of photo i, evaluated at the initial approximations. Matrix $\ddot{B}_{ij}$ contains the partial derivatives of the collinearity equations with respect to the object space coordinates of point j, evaluated at the initial approximations. Matrix $\dot{\Delta}_i$ contains corrections for the initial approximations of the exterior orientation parameters for photo i, and matrix $\ddot{\Delta}_j$ contains corrections for the initial approximations of the object space coordinates of point j. Matrix ε_{ij} contains measured minus computed x and y photo coordinates for point j on photo i, and finally matrix V_{ij} contains residuals for the x and y photo coordinates.

Proper weights must be assigned to photo coordinate observations in order to be included in the bundle adjustment. Expressed in matrix form, the weights for x and y photo coordinate observations of point j on photo i are

$$W_{ij} = \sigma_o^2 \begin{bmatrix} \sigma_{x_{ij}}^2 & \sigma_{x_{ij}y_{ij}} \\ \sigma_{y_{ij}x_{ij}} & \sigma_{y_{ij}}^2 \end{bmatrix}^{-1} \tag{17-5}$$

where σ_o^2 is the reference variance; $\sigma_{x_{ij}}^2$ and $\sigma_{y_{ij}}^2$ are variances in x_{ij} and y_{ij}, respectively; and

$\sigma_{x_{ij}y_{ij}} = \sigma_{y_{ij}x_{ij}}$ is the covariance of x_{ij} with y_{ij}. The reference variance is an arbitrary parameter which can be set equal to 1, and in many cases, the covariance in photo coordinates is equal to zero. In this case, the weight matrix for photo coordinates simplifies to

$$W_{ij} = \begin{bmatrix} \frac{1}{\sigma^2_{x_{ij}}} & 0 \\ 0 & \frac{1}{\sigma^2_{y_{ij}}} \end{bmatrix} \tag{17-6}$$

The next type of observation to be considered is ground control. Observation equations for ground control coordinates are

$$\begin{aligned} X_j &= X_j^{00} + v_{X_j} \\ Y_j &= Y_j^{00} + v_{Y_j} \\ Z_j &= Z_j^{00} + v_{Z_j} \end{aligned} \tag{17-7}$$

where X_j, Y_j, and Z_j are unknown coordinates of point j; X_j^{00}, Y_j^{00}, and Z_j^{00} are the measured coordinate values for point j; and v_{X_j}, v_{Y_j}, and v_{Z_j} are the coordinate residuals for point j.

Even though ground control observation equations are linear, in order to be consistent with the collinearity equations, they will also be approximated by the first-order terms of Taylor's series.

$$\begin{aligned} X_j^0 + dX_j &= X_j^{00} + v_{X_j} \\ Y_j^0 + dY_j &= Y_j^{00} + v_{Y_j} \\ Z_j^0 + dZ_j &= Z_j^{00} + v_{Z_j} \end{aligned} \tag{17-8}$$

In Eq. (17-8), X_j^0, Y_j^0, and Z_j^0 are initial approximations for the coordinates of point j; dX_j, dY_j, and dZ_j are corrections to the approximations for the coordinates of point j; and the other terms are as previously defined.

Rearranging the terms of Eq. (17-8) and expressing the result in matrix form gives

$$\ddot{\Delta}_j = \ddot{C}_j + \ddot{V}_j \tag{17-9}$$

where $\ddot{\Delta}_j$ is as previously defined and

$$\ddot{C}_j = \begin{bmatrix} X_j^{00} - X_j^0 \\ Y_j^{00} - Y_j^0 \\ Z_j^{00} - Z_j^0 \end{bmatrix} \qquad \ddot{V}_j = \begin{bmatrix} v_{X_j} \\ v_{Y_j} \\ v_{Z_j} \end{bmatrix}$$

As with photo coordinate measurements, proper weights must be assigned to ground control

coordinate observations in order to be included in the bundle adjustment. Expressed in matrix form, the weights for *X*, *Y*, and *Z* ground control coordinate observations of point *j* are

$$\ddot{W}_j = \sigma_o^2 \begin{bmatrix} \sigma_{X_j}^2 & \sigma_{X_jY_j} & \sigma_{X_jZ_j} \\ \sigma_{Y_jX_j} & \sigma_{Y_j}^2 & \sigma_{Y_jZ_j} \\ \sigma_{Z_jX_j} & \sigma_{Z_jY_j} & \sigma_{Z_j}^2 \end{bmatrix}^{-1} \tag{17-10}$$

where σ_o^2 is the reference variance; $\sigma_{X_j}^2$, $\sigma_{Y_j}^2$, and $\sigma_{Z_j}^2$ are the variances in X_j^{00}, Y_j^{00}, and Z_j^{00}, respectively; $\sigma_{X_jY_j} = \sigma_{Y_jX_j}$ is the covariance of X_j^{00} with Y_j^{00}; $\sigma_{X_jZ_j} = \sigma_{Z_jX_j}$ is the covariance of X_j^{00} with Z_j^{00}; and $\sigma_{Y_jZ_j} = \sigma_{Z_jY_j}$ is the covariance of Y_j^{00} with Z_j^{00}. As before, the reference variance can be arbitrarily set equal to 1; however, in general, since ground control coordinates are indirectly determined quantities, their covariances are not equal to zero.

The final type of observation consists of measurements of exterior orientation parameters. The form of their observation equations is similar to that of ground control and given as Eq. (17-11).

$$\begin{aligned} \omega_i &= \omega_i^{00} + v_{\omega_i} \\ \phi_i &= \phi_i^{00} + v_{\phi_i} \\ \kappa_i &= \kappa_i^{00} + v_{\kappa_i} \\ X_{L_i} &= X_{L_i}^{00} + v_{X_{L_i}} \\ Y_{L_i} &= Y_{L_i}^{00} + v_{Y_{L_i}} \\ Z_{L_i} &= Z_{L_i}^{00} + v_{Z_{L_i}} \end{aligned} \tag{17-11}$$

$$\Delta_i = C_i + V_i \tag{17-12}$$

The weight matrix for exterior orientation parameters has the following form:

$$\dot{W}_i = \begin{bmatrix} \sigma_{\omega_i}^2 & \sigma_{\omega_i\phi_i} & \sigma_{\omega_i\kappa_i} & \sigma_{\omega_iX_{L_i}} & \sigma_{\omega_iY_{L_i}} & \sigma_{\omega_iZ_{L_i}} \\ \sigma_{\phi_i\omega_i} & \sigma_{\phi_i}^2 & \sigma_{\phi_i\kappa_i} & \sigma_{\phi_iX_{L_i}} & \sigma_{\phi_iY_{L_i}} & \sigma_{\phi_iZ_{L_i}} \\ \sigma_{\kappa_i\omega_i} & \sigma_{\kappa_i\phi_i} & \sigma_{\kappa_i}^2 & \sigma_{\kappa_iX_{L_i}} & \sigma_{\kappa_iX_{L_i}} & \sigma_{\kappa_iZ_{L_i}} \\ \sigma_{X_{L_i}\omega_i} & \sigma_{X_{L_i}\phi_i} & \sigma_{X_{L_i}\kappa_i} & \sigma_{X_{L_i}}^2 & \sigma_{X_{L_i}Y_{L_i}} & \sigma_{X_{L_i}Z_{L_i}} \\ \sigma_{Y_{L_i}\omega_i} & \sigma_{Y_{L_i}\phi_i} & \sigma_{Y_{L_i}\kappa_i} & \sigma_{Y_{L_i}X_{L_i}} & \sigma_{Y_{L_i}}^2 & \sigma_{Y_{L_i}Z_{L_i}} \\ \sigma_{Z_{L_i}\omega_i} & \sigma_{Z_{L_i}\phi_i} & \sigma_{Z_{L_i}\kappa_i} & \sigma_{Z_{L_i}X_{L_i}} & \sigma_{Z_{L_i}Y_{L_i}} & \sigma_{Z_{L_i}}^2 \end{bmatrix}^{-1} \tag{17-13}$$

With the observation equations and weights defined as above, the full set of normal equations may be formed directly. In matrix form, the full normal equations are

$$N\Delta = K$$

where

$$N = \begin{bmatrix}
\dot{N}_1 + \dot{W}_1 & {}_6\dot{0}^6 & {}_6\dot{0}^6 & \cdots & {}_6\dot{0}^6 & \bar{N}_{11} & \bar{N}_{12} & \bar{N}_{13} & \cdots & \bar{N}_{1n} \\
{}_6\dot{0}^6 & \dot{N}_2 + \dot{W}_2 & {}_6\dot{0}^6 & \cdots & {}_6\dot{0}^6 & \bar{N}_{21} & \bar{N}_{22} & \bar{N}_{23} & \cdots & \bar{N}_{2n} \\
{}_6\dot{0}^6 & {}_6\dot{0}^6 & \dot{N}_3 + \dot{W}_3 & \cdots & {}_6\dot{0}^6 & \bar{N}_{31} & \bar{N}_{32} & \bar{N}_{33} & \cdots & \bar{N}_{3n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
{}_6\dot{0}^6 & {}_6\dot{0}^6 & {}_6\dot{0}^6 & \cdots & \dot{N}_m + \dot{W}_m & \bar{N}_{m1} & \bar{N}_{m2} & \bar{N}_{m3} & \cdots & \bar{N}_{mn} \\
\bar{N}_{11}^T & \bar{N}_{21}^T & \bar{N}_{31}^T & \cdots & \bar{N}_{m1}^T & \ddot{N}_1 + \ddot{W}_1 & {}_3\dot{0}^3 & {}_3\dot{0}^3 & \cdots & {}_3\dot{0}^3 \\
\bar{N}_{12}^T & \bar{N}_{22}^T & \bar{N}_{32}^T & \cdots & \bar{N}_{m2}^T & {}_3\dot{0}^3 & \ddot{N}_2 + \ddot{W}_2 & {}_3\dot{0}^3 & \cdots & {}_3\dot{0}^3 \\
\bar{N}_{13}^T & \bar{N}_{23}^T & \bar{N}_{33}^T & \cdots & \bar{N}_{m3}^T & {}_3\dot{0}^3 & {}_3\dot{0}^3 & \ddot{N}_3 + \ddot{W}_3 & \cdots & {}_3\dot{0}^3 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\bar{N}_{1n}^T & \bar{N}_{2n}^T & \bar{N}_{3n}^T & \cdots & \bar{N}_{mn}^T & {}_3\dot{0}^3 & {}_3\dot{0}^3 & {}_3\dot{0}^3 & \cdots & \ddot{N}_n + \ddot{W}_n
\end{bmatrix}$$

$$\Delta = \begin{bmatrix} \dot{\Delta}_1 \\ \dot{\Delta}_2 \\ \dot{\Delta}_3 \\ \vdots \\ \dot{\Delta}_m \\ \ddot{\Delta}_1 \\ \ddot{\Delta}_2 \\ \ddot{\Delta}_3 \\ \vdots \\ \ddot{\Delta}_n \end{bmatrix} \qquad K = \begin{bmatrix} \dot{K}_1 + \dot{W}_1\dot{C}_1 \\ \dot{K}_2 + \dot{W}_1\dot{C}_2 \\ \dot{K}_3 + \dot{W}_1\dot{C}_3 \\ \vdots \\ \dot{K}_m + \dot{W}_m\dot{C}_m \\ \ddot{K}_1 + \ddot{W}_1\ddot{C}_1 \\ \ddot{K}_2 + \ddot{W}_2\ddot{C}_2 \\ \ddot{K}_3 + \ddot{W}_3\ddot{C}_3 \\ \vdots \\ \ddot{K}_n + \ddot{W}_n\ddot{C}_n \end{bmatrix}$$

The submatrices in the above forms are

$$\dot{N}_i = \sum_{j=1}^{n} \dot{B}_{ij}^T W_{ij} \dot{B}_{ij} \qquad \bar{N}_{ij} = \dot{B}_{ij}^T W_{ij} \ddot{B}_{ij} \qquad \ddot{N}_j = \sum_{i=1}^{m} \ddot{B}_{ij}^T W_{ij} \ddot{B}_{ij}$$

$$\dot{K}_i = \sum_{j=1}^{n} \dot{B}_{ij}^T W_{ij} \varepsilon_{ij} \qquad \ddot{K}_j = \sum_{i=1}^{m} \ddot{B}_{ij}^T W_{ij} \varepsilon_{ij}$$

In the above expressions, *m* is the number of photos, *n* is the number of points, *i* is the photo subscript, and *j* is the point subscript. Note that if point *j* does not appear on photo *i*, the corresponding submatrix will be a zero matrix. Note also that the $\dot{W}_i$ contributions to the *N* matrix and the $\dot{W}_i\dot{C}_i$ contributions to the *K* matrix are made only when observations for exterior orientation parameters exist; and the $\ddot{W}_j$ contributions to the *N* matrix and the $\ddot{W}_j\ddot{C}_j$ contributions to the *K* matrix are made only

for ground control point observations.

While the normal equations are being formed, it is recommended that the estimate for the standard deviation of unit weight be calculated (see Sec. B-10). Assuming the initial approximations are reasonable, matrices ε_{ij}, $\dot{C}_i$, and $\ddot{C}_j$ are good estimates of the negatives of the residuals. Therefore, the estimate of the standard deviation of unit weight can be computed by

$$S_0 = \sqrt{\frac{\sum_{j=1}^{n}\sum_{i=1}^{m}\varepsilon_{ij}^T W_{ij}\varepsilon_{ij} + \sum_{i=1}^{m}\dot{C}_i^T \dot{W}_i \dot{C}_i + \sum_{j=1}^{n}\ddot{C}_j^T \ddot{W}_j \ddot{C}_j}{\text{n.o.} - \text{n.u.}}} \tag{17-15}$$

In Eq. (17-15), n.o. is the number of observations and n.u. is the number of unknowns in the solution. If all observations have been properly weighted, S_0 should be close to 1.

After the normal equations have been formed, they are solved for the unknowns Δ, which are corrections to the initial approximations for exterior orientation parameters and object space coordinates. The corrections are then added to the approximations, and the procedure is repeated until the estimated standard deviation of unit weight converges. At that point, the covariance matrix for the unknowns can be computed by

$$\Sigma_{\Delta\Delta} = S_0^2 N^{-1} \tag{17-16}$$

Computed standard deviations for the unknowns can then be obtained by taking the square root of the diagonal elements of the $\Sigma_{\Delta\Delta}$ matrix.

17-7 Initial Approximations for the Bundle Adjustment

In Sec. 17-6, the equations involved in a bundle adjustment were presented. Since the equations are nonlinear, Taylor's series was used to linearize the equations; therefore, initial approximations are required for the unknowns. For the solution to converge properly, the initial approximations must be reasonably close to the correct values. Several methods may be used to obtain initial approximations; however, preliminary strip adjustments are most commonly employed.

The first step is to perform analytical relative orientation (see Sec. 11-10) for each stereopair in the block. The photo coordinate residuals should be inspected at this point as an initial check on the measurements. Next, the relatively oriented models are connected to form strips by the method presented in Sec. 17-4. Residuals from this step can also provide a quality check on the photo coordinate measurements and point identification. After all the strip models have been formed and validated, each strip is individually adjusted to ground control points located within each strip. This adjustment to ground control can be performed either by a three-dimensional conformal coordinate transformation or by a three-dimensional polynomial transformation (see Sec. 17-5). Residuals from this step provide a check on the ground control coordinates as well as point identification.

At this point, ground coordinates will have been calculated for all points in the photo block. An additional check can be performed to validate the identification of tie points between strips. If the identification of tie points is consistent, their coordinates as determined in adjacent strips should agree within a small tolerance. Assuming everything is consistent at this point, the resulting ground coordinates can be used as initial approximations for the bundle adjustment.

Approximations for the exterior orientation parameters can also be obtained directly from the strip adjustment if the adjustment is performed using a three-dimensional conformal coordinate transformation. In that case, since perspective centers (camera stations) are included when adjacent models are connected, their object space coordinates will be available after the final adjustment to ground control. Assuming vertical photography, zeros can be used as approximations for ω and ϕ. Approximations for κ can be obtained directly from the final three-dimensional conformal coordinate transformation to ground control, which contains a compatible κ angle. If a polynomial strip adjustment is performed, the perspective centers are not included in the adjustment. In that case, after the polynomial adjustment is completed, the space resection problem can be solved for each photo (see Sec. 11-6). In these calculations, the ground coordinates obtained for the pass points in the polynomial adjustment are used as control coordinates.

In cases where more precise initial approximations are needed, one can "chain" together the rotations of a full strip through relative orientation to obtain estimates of ω, φ, and κ. Precise initial approximations decrease the number of iterations required for convergence, and can therefore significantly increase the speed of bundle adjustment solutions. The final three-dimensional conformal coordinate transformation from sequential independent model triangulation (see Sec. 17-5) provides the approximations of orientation angles for the first photo in the strip. Next, a rotation matrix from ground to strip, M_{3D}, is formed from these angles, and a rotation matrix from the first photo to the second photo, M_{1-2}, is formed from the rotation angles of the first relative orientation via the definitions in Sec. C-7. The product of these matrices $M_{1-2}M_{3D}$, yields the rotation matrix from ground to the second photo and can therefore be used to obtain approximations of ω_2, φ_2, and κ_2 for the second photo. Approximations for all other photos in the strip can be obtained by repeating this process (i.e., *chaining* the angular orientations together).

Example 17-2

A strip model was constructed sequentially using the method described in Sec. 17-4, and then adjusted to ground using a three-dimensional conformal coordinate transformation. Using the data provided in the table, find the initial approximations of ω, φ, and κ for each photo in the strip using the chain method.

	ω	φ	κ
Relative orientation of photo 1 to photo 2	0.7881°	2.3853°	0.7254°
Relative orientation of photo 2 to photo 3	0.0685°	−1.3379°	1.7563°
Relative orientation of photo 3 to photo 4	0.4679°	−1.9940°	−2.2502°
Orientation of ground to strip	2.1566°	1.3424°	176.7343°

Solution Use the rotation angles for the orientation of ground to strip to approximate ω, φ, and κ for the first photo. Form the rotation matrix from the ground to strip system using the definitions in Sec. C-7:

$$M_{3D} = \begin{bmatrix} -0.9981 & 0.0560 & 0.0255 \\ -0.0570 & -0.9977 & -0.0362 \\ 0.0234 & -0.0376 & 0.9990 \end{bmatrix}$$

Next, form the rotation matrix representing the relative angular orientation from photo 1 to photo 2:

$$M_{1\text{-}2} = \begin{bmatrix} 0.9991 & 0.0132 & -0.0414 \\ -0.0126 & 0.9998 & 0.0143 \\ 0.0416 & -0.0137 & 0.9990 \end{bmatrix}$$

The product of these matrices yields the rotation matrix from ground to photo 2 which can be used to find approximate values for ω, φ, and κ for photo 2 via the method described in Sec. D-10:

$$M_{1\text{-}2}M_{3D} = \begin{bmatrix} 0.9989 & 0.0444 & -0.0164 \\ -0.0440 & -0.9998 & -0.0223 \\ -0.0174 & -0.0215 & 0.9996 \end{bmatrix}$$

	ω	φ	κ
Orientation of ground to photo 2	1.2345°	−0.9943°	177.4789°

Multiplying the rotation matrix formed by the relative orientation angles from photo 2 to photo 3, $M_{2\text{-}3}$, by the above matrix yields the rotation matrix from ground to photo 3 which can be used to find approximate values for ω, φ, and κ for photo 3:

$$M_{2\text{-}3}(M_{1\text{-}2}M_{3D}) = \begin{bmatrix} 0.9993 & 0.0306 & 0.0234 \\ -0.0306 & 0.9995 & 0.0005 \\ -0.0233 & -0.0012 & 0.9997 \end{bmatrix} \begin{bmatrix} 0.9989 & 0.0444 & -0.0164 \\ -0.0440 & -0.9998 & -0.0223 \\ -0.0174 & -0.0215 & 0.9996 \end{bmatrix}$$

$$= \begin{bmatrix} -0.9999 & 0.0132 & 0.0063 \\ -0.0134 & -0.9997 & -0.0213 \\ -0.0060 & -0.0214 & 0.9998 \end{bmatrix}$$

	ω	φ	κ
Orientation of ground to photo 3	1.2249°	0.3453°	179.2343°

Multiplying the rotation matrix formed by the relative orientation angles from photo 3 to photo 4, $M_{3\text{-}4}$, by the above matrix yields the rotation matrix from ground to photo 4 which can be used to find approximate values for ω, φ, and κ for photo 4:

$$M_{3\text{-}4}(M_{2\text{-}3}M_{1\text{-}2}M_{3D}) = \begin{bmatrix} 0.9986 & -0.0396 & 0.0344 \\ 0.0392 & 0.9992 & 0.0095 \\ -0.0348 & -0.0082 & 0.9994 \end{bmatrix} \begin{bmatrix} -0.9999 & 0.0132 & 0.0063 \\ -0.0134 & -0.9997 & -0.0213 \\ -0.0060 & -0.0214 & 0.9998 \end{bmatrix}$$

$$= \begin{bmatrix} -0.9978 & 0.0520 & 0.0416 \\ -0.0525 & -0.9986 & -0.0115 \\ 0.0409 & -0.0137 & 0.9991 \end{bmatrix}$$

	ω	φ	κ
Orientation of ground to photo 4	0.7836°	2.3454°	176.9863°

▲

17-8 Bundle Adjustment with Airborne GPS Control

As mentioned in Sec. 17-1, kinematic GPS and INS observations can be taken aboard the aircraft as the photography is being acquired to determine coordinates and angular attitude for exposure stations. Use of GPS and INS in the aircraft to control a bundle adjustment of a block of photographs is termed *airborne control*. By including coordinates of the exposure stations and angular attitude of the camera in the adjustment, the amount of ground control can be greatly reduced.

Figure 17-6 illustrates the geometric relationship between a camera, inertial measurement unit (IMU), and GPS antenna on an aircraft. In this figure, x, y, and z represent the standard three-dimensional coordinate system of a mapping camera; and x_A, y_A, and z_A represent the coordinates of the GPS antenna relative to the camera axes, often referred to as the lever arm. The x axis of the camera is parallel to the longitudinal axis of the aircraft, the z axis is vertical, and the y axis is perpendicular to the x and z axes. Since object space coordinates obtained by GPS pertain to the phase center of the antenna but the exposure station is defined as the incident nodal point of the camera lens, the GPS coordinates of the antenna must be translated to the camera lens. To properly compute the translations, it is necessary to know the angular orientation of the camera with respect to the object space coordinate system. Determining the correct angular orientation is complicated by the use of a gimbaled camera mount which allows relative rotations between the camera and the aircraft frame.

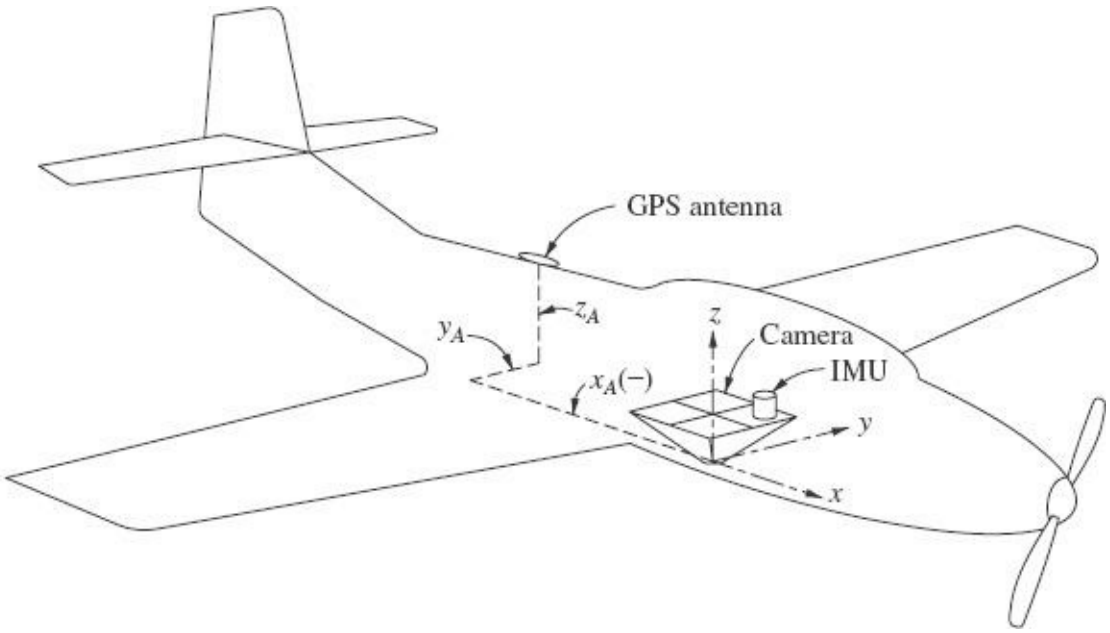


FIGURE 17-6 Configuration of camera, IMU, and GPS antenna for airborne GPS control.

If the camera in its mount was fixed, the rotation matrix M_i, consisting of angular orientation parameters of the camera (ω_i, ϕ_i and κ_i) would translate directly to angular orientation of the camera-to-antenna vector. However, differential rotation from the airframe to the camera, represented by M_i^m (the superscript m stands for *mount*), must also be taken into account in order to determine the angular attitude of the camera-to-antenna vector in object space. Note that even in a so-called fixed mount there will generally be a crab adjustment, rotation about the z axis of the fixed-mount coordinate system, to ensure proper photographic coverage (see Sec. 3-6). Some camera mounts such as the Leica PAV30 shown in Fig. 3-8 have the capability of measuring the differential rotations, and they can be recorded by a computer. The following equation specifies the rotation of the camera-to-antenna vector with respect to object space:

$$M_i' = M_i^m M_i \tag{17-17}$$

In Eq. (17-17), M_i is the conventional rotation matrix consisting of angular exterior orientation parameters of the camera with respect to the object space coordinate system (ω_i, ϕ_i and κ_i); M_i^m in the rotation matrix of the camera with respect to the mount; and M_i' is the rotation matrix of the camera-to-antenna vector with respect to object space coordinates.

Once M_i' has been determined, the rotation angles (ω_i', ϕ_i', and κ_i') can be computed. [See Eqs. (C-33) and Sec. D-10.] After M_i' has been computed, the coordinates of the camera lens can be computed by Eq. (17-18). (Note: subscript *i* has been dropped.)

$$\begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix} = \begin{bmatrix} X_{GPS} \\ Y_{GPS} \\ Z_{GPS} \end{bmatrix} - M'^T \begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix} \tag{17-18}$$

When a camera mount is used which does not provide for measurement of the differential rotation from the airframe to the camera, it is assumed to be equal to zero, resulting in errors in the computed position of the camera lens. This error can be minimized by mounting the GPS antenna vertically above the camera in the aircraft, which effectively eliminates the error due to unaccounted crab adjustment, rotation about the *z* axis of the fixed-mount coordinate system. As long as the differential tilt rotations are small (less than a couple of degrees) and the antenna-to-camera vector is short (less than 2 m), the lens positional error will be less than 10 cm. One last comment must be made concerning the translation of GPS antenna coordinates to the lens. Since the values of ω, ϕ, and κ are required to compute the translation, the antenna offset correction must be included within the iterative loop of the analytical bundle adjustment.

In order to use airborne control, it is necessary to have accurate values for the lever arm between the camera and GPS antenna. Perhaps the most common method for determining this vector is direct measurement using conventional surveying techniques. However, it is possible to include their values as unknowns in the bundle adjustment solution. Equations (17-19) show the collinearity equations for an imaged point with observations from GPS and lever arm parameters included. Note that the lever arm parameters are included under the assumption that the camera mount is fixed and should otherwise reflect the change in angular attitude due to rotation of the mount.

$$x_{ij} = x_0 - f\left[\frac{m_{11i}(X_j - X_{\text{GPS}}) + m_{12i}(Y_j - Y_{\text{GPS}}) + m_{13i}(Z_j - Z_{\text{GPS}}) + x_A}{m_{31i}(X_j - X_{\text{GPS}}) + m_{32i}(Y_j - Y_{\text{GPS}}) + m_{33i}(Z_j - Z_{\text{GPS}}) + z_A}\right]$$

$$y_{ij} = y_0 - f\left[\frac{m_{21i}(X_j - X_{\text{GPS}}) + m_{22i}(Y_j - Y_{\text{GPS}}) + m_{23i}(Z_j - Z_{\text{GPS}}) + y_A}{m_{31i}(X_j - X_{\text{GPS}}) + m_{32i}(Y_j - Y_{\text{GPS}}) + m_{33i}(Z_j - Z_{\text{GPS}}) + z_A}\right] \tag{17-19}$$

The lever arm parameters are highly correlated with both the interior and exterior orientation parameters. This can greatly affect the precision of their solution in the bundle adjustment, which is why the lever arm parameters are normally measured using conventional surveying techniques.

The boresight angles that define the orientation of the IMU with respect to the camera can be found by the difference between results from a bundle adjustment using ground control, and the values obtained from using airborne control. Alternatively, as with the lever arm parameters, these values can also be included in a bundle adjustment as unknowns. In this case, the *m*s in Eq. (17-19) correspond to

matrix entries of M_i, the product of the rotation matrix determined by the INS, M_i^{IMU}, and the boresight rotation matrix, ΔM as shown in Eq. (17-20).

$$M_i = \Delta M M_i^{\text{IMU}} \tag{17-20}$$

Since two rotation matrices are included, there are six unknown rotation angles in Eq. (17-19). This makes the linearization of the collinearity equations significantly more complex than with the standard formulation.

Another consideration regarding airborne GPS positioning is the problem of *loss of lock* on the GPS satellites, especially during banked turns. When a GPS receiver operating in the kinematic mode loses lock on too many satellites, the integer ambiguities must be redetermined (see Chap. 16-8). Since returning to a previously surveyed point is generally out of the question, on-the-fly (OTF) techniques are used to calculate the correct integer ambiguities. With high-quality, dual-frequency, P-code receivers, OTF techniques are often successful in correctly redetermining the integer ambiguities. In some cases, however, an integer ambiguity solution may be obtained which is slightly incorrect. This results in an approximately linear drift in position along the flight line, which causes exposure station coordinate errors to deteriorate. This problem can be detected by using a small number of ground control points at the edges of the photo block. Inclusion of additional parameters in the adjustment corresponding to the linear drift enables a correction to be applied which eliminates this source of error. Often, *cross strips* are flown at the ends of the regular block strips, as shown in Fig. 17-7. The cross strips contain ground control points at each end which allow drift due to incorrect OTF integer ambiguities to be detected and corrected. The corrected cross strips in turn serve to provide endpoint coordinates for the remainder of the strips in the block, thus enabling drift corrections to be made for those strips as well.

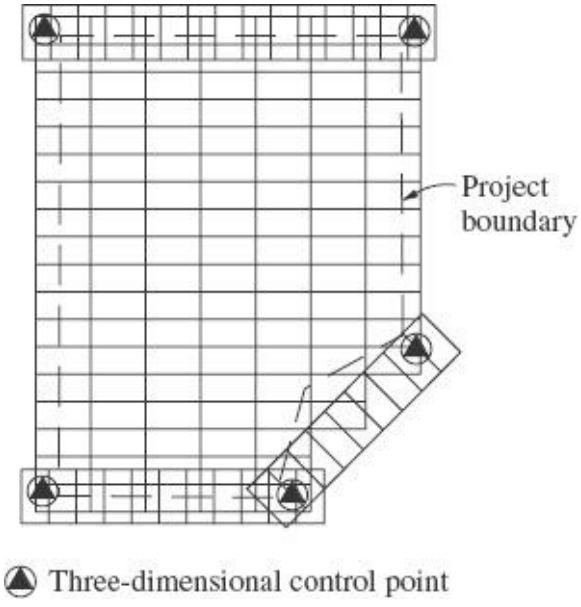


FIGURE 17-7 Configuration of flight strips for airborne GPS control.

Two additional precautions regarding airborne GPS should be noted. First, it is recommended that a bundle adjustment with analytical self-calibration (see Sec. 19-4) be employed when airborne GPS control is used. Often, due to inadequate modeling of atmospheric refraction distortion, strict enforcement of the calibrated principal distance (focal length) of the camera will cause distortions and

excessive residuals in photo coordinates. Use of analytical self-calibration will essentially eliminate that effect. Second, it is essential that appropriate object space coordinate systems be employed in data reduction. GPS coordinates in a geocentric coordinate system should be converted to local vertical coordinates for the adjustment (see Secs. 5-5 and F-4). After aerotriangulation is completed, the local vertical coordinates can be converted to whatever system is desired. Elevations relative to the ellipsoid can be converted to orthometric elevations by using an appropriate geoid model.

17-9 Interpretation of Bundle Adjustment Results

After setting up and executing a bundle adjustment, it is essential to analyze and interpret the results. Analysis of the results allows one to identify blunders, maximize the precision of the solution, and report the level of confidence in the solved parameters. The output of a bundle adjustment program can aid in troubleshooting divergent or grossly inaccurate solutions. Similarly, scrutiny of residual statistics provides a guide for the user to tune a priori estimates of observation standard deviations, and increase the precision of seemingly sufficient solutions. Fortunately, the least squares methods described at the end of Sec. 17-6 and in Sec. B-10, provide tools that can be used toward these goals, namely the observation residuals and the covariance matrix. Problems with bundle adjustments come in many forms. These include divergence, convergence with blundered observations, and convergence with improper weighting of observations.

During execution, a bundle adjustment may fail to converge on a solution. Divergence occurs when the standard deviation of unit weight, S_0, gets larger at successive iterations. A typical bundle adjustment program will recognize divergence and terminate when this happens. If a bundle adjustment diverges, there are a few likely causes. The first is that the initial approximations were not close enough to their optimal values in a least squares sense. This is usually the result of blunders in reporting the initial approximations such as reversed direction of κ, or incorrect geometric configuration of the object space coordinates of points (e.g., point distribution was inverted). Although some software may detect it, the inclusion of a pass point with coordinates for only one photo can cause a bundle adjustment to diverge. A related blunder is misidentification of a point on a photo, which can easily occur if multiple points are associated with similar-looking features such as manholes. The best way to identify blunders in initial approximations is to simply recheck their values and make sure they are logical with respect to the geometry of the photo block. For example, if the plane was flying east to west when a photo was taken, then κ should be close to 180°. In the same scenario, if a point a is imaged farther along the x axis than point b, then the ground coordinates for point b will have higher X coordinates (easting) than point a (recall that the x direction in an image is assumed to be nominally aligned with the direction of flight). The sequential approach to semianalytical aerotriangulation presented in Secs. 17-4 and 17-5, in addition to providing initial approximations, can be used as quality control prior to the bundle adjustment. The sequential approach provides a check in the form of pass point and control point residuals from the relative orientation between images, transformations between models, and the transformation from strip to ground.

In some cases, a bundle adjustment may converge even with blundered observations present. There are *robust* methods for least squares adjustments (such as Random Sample Consensus, RANSAC) that allow for the automatic identification and elimination of blunders. However, in many circumstances, blunders can be identified using simpler methods or manually. The first indication that a converged bundle adjustment contains one or more blunders is an inordinately high S_0. Blundered observations will have relatively large residuals. A first check is to look for residuals significantly

greater than the estimated a priori standard deviation. Similarly, since the residuals are assumed to be normally distributed, one should be suspicious of any observation with a residual greater than 3 times its standard deviation. When a blunder is found, the best course of action is to remeasure the point. However, if there is sufficient point coverage near the point, then it may be removed. In some cases a point may seem to be a blunder, but is actually just measured less accurately than other points. This can be caused by the point being in an area on the photo that is less conducive to point matching, like a field of grass. When this happens, the best strategy is to "loosen" the point in the adjustment, i.e., increase the a priori standard deviation of the observation relative to other observations. Once corrections and tuning have been applied to the input, the adjustment should be run again. Sometimes multiple bundle adjustments for a single data set must be executed in order to eliminate all blunders and to fine tune a priori precision estimates. Between each execution, it is sometimes helpful to update initial approximations using the previous adjustment allowing for faster convergence.

In the best case scenario, the adjustment will converge on a solution and have a standard deviation of unit weight close to one. This indicates that initial approximations for unknown parameters were sufficient and that a priori standard deviations reflected the true values of the precision of the observations. It may occur that a bundle adjustment converges and has no blunders but S_0 is either too high or too low. In general, this will not significantly affect the results, but it is still suggested that it be corrected. The reason this occurs is that a priori estimates of standard deviations are incorrect. If S_0 is higher than one, one should increase the a priori standard deviations. If S_0 is lower than one, the a priori standard deviations should be decreased.

Post-bundle adjustment statistics provide an excellent measure of the quality of the solution. After a bundle adjustment converges properly, the standard deviations for both exterior orientation parameters and ground coordinates of pass points are obtained using Eq. (17-16). These a posteriori standard deviations can be used to quantify the precision of the solution. The geometry of the adjustment (the relative location of photos, pass points, and control points), the precision of the observations, and the redundancy of the observations influence the standard deviations. For example, the exterior orientation parameter φ for a photo in an east-west strip such as the one illustrated in Fig. 17-4 will typically have a lower standard deviation than that for ω. Due to the geometry of the strip, small changes in the position of points (e.g., errors) would influence the solved rotation about the x axis more than the rotation about the y axis. It is common for the solved horizontal ground coordinates of pass points to have higher standard deviations than Z coordinates. One way to visualize why this occurs is to imagine rays from the perspective centers of two near-vertical photos intersecting at a point on the ground. Small changes in either the position of the point in the image or the exterior orientation parameters of the camera will affect the point of closest intersection of the rays more in the Z direction than in the X or Y directions. Another geometric factor to consider is the location of the point in the photo. For example, and for a similar reason as above, points imaged farther from the x axis in an east-west strip of photos will have higher standard deviations for Y ground coordinates than those with a more central location in the photo. In addition to geometry, redundancy of observations can have a large influence on the precision of the bundle adjustment solution. For instance, points measured in tri-lap areas can be expected to have smaller standard deviations than points only imaged in two photos.

It should be noted that although a posteriori standard deviations can sufficiently represent the quality of adjustments, coordinate comparisons with *checkpoints*—precisely measured control points that are not included in the adjustment—are generally better measures of the accuracy of the solution. A drawback of using checkpoints is that they require extra effort in the form of field work and/or photogrammetric procedures. However, their coordinates can be obtained any time after photo acquisition and adjustment as long as they are distinguishable in the images can still be physically

located on the ground.

17-10 Aerotriangulation with Airborne Linear Array Sensors

Linear array sensors capture images with different geometry compared to the point perspective of a frame camera. Each scan line has its own set of exterior orientation parameters compared to one set of parameters for a frame camera image. The aircraft must be equipped with a GPS-INS system in order to measure these exterior orientation parameters for each scan line of the image. Aerial three-line linear array sensors, such as the Leica ADS80, have the advantage of providing three different perspectives of ground points along a strip from collection of forward, backward, and nadir scans. This facilitates the use of aerotriangulation toward improving the accuracy of the data. Figure 17-8 illustrates the geometry of a three-line linear array system.

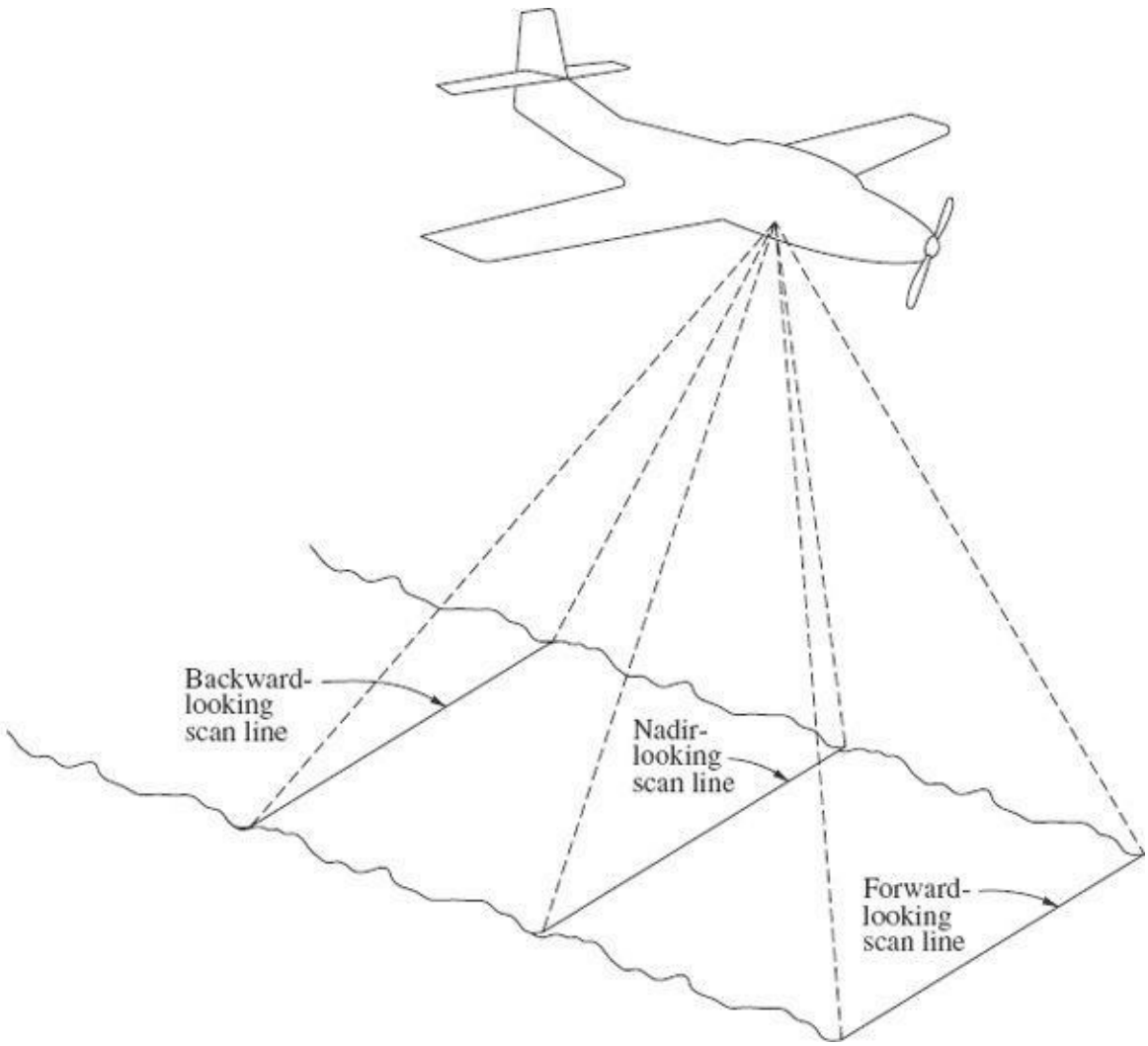


FIGURE 17-8 Three-line linear array sensor scans: forward, nadir, and backward.

Three-line scanners collect three raw image scenes synchronously along a strip. One scene consists of the collection of scan lines from the backward-looking linear array, another is from the nadir-looking linear array, and the third is from the forward-looking linear array. In their raw format, Level 0, these scenes are distorted due to aircraft movement during collection. Correcting the data for sensor tilt and aircraft movement using GPS-INS measurements yields nominally rectified imagery, Level 1. Figure 17-9 shows Level 0 imagery and Level 1 imagery. In the ADS systems, the

transformations from Level 0 to Level 1 are done in real time. In order to increase the accuracy of the imagery and to facilitate the calibration of boresight and lever arm parameters, the exterior orientation parameters obtained by GPS-INS are adjusted using a unique method of aerotriangulation.

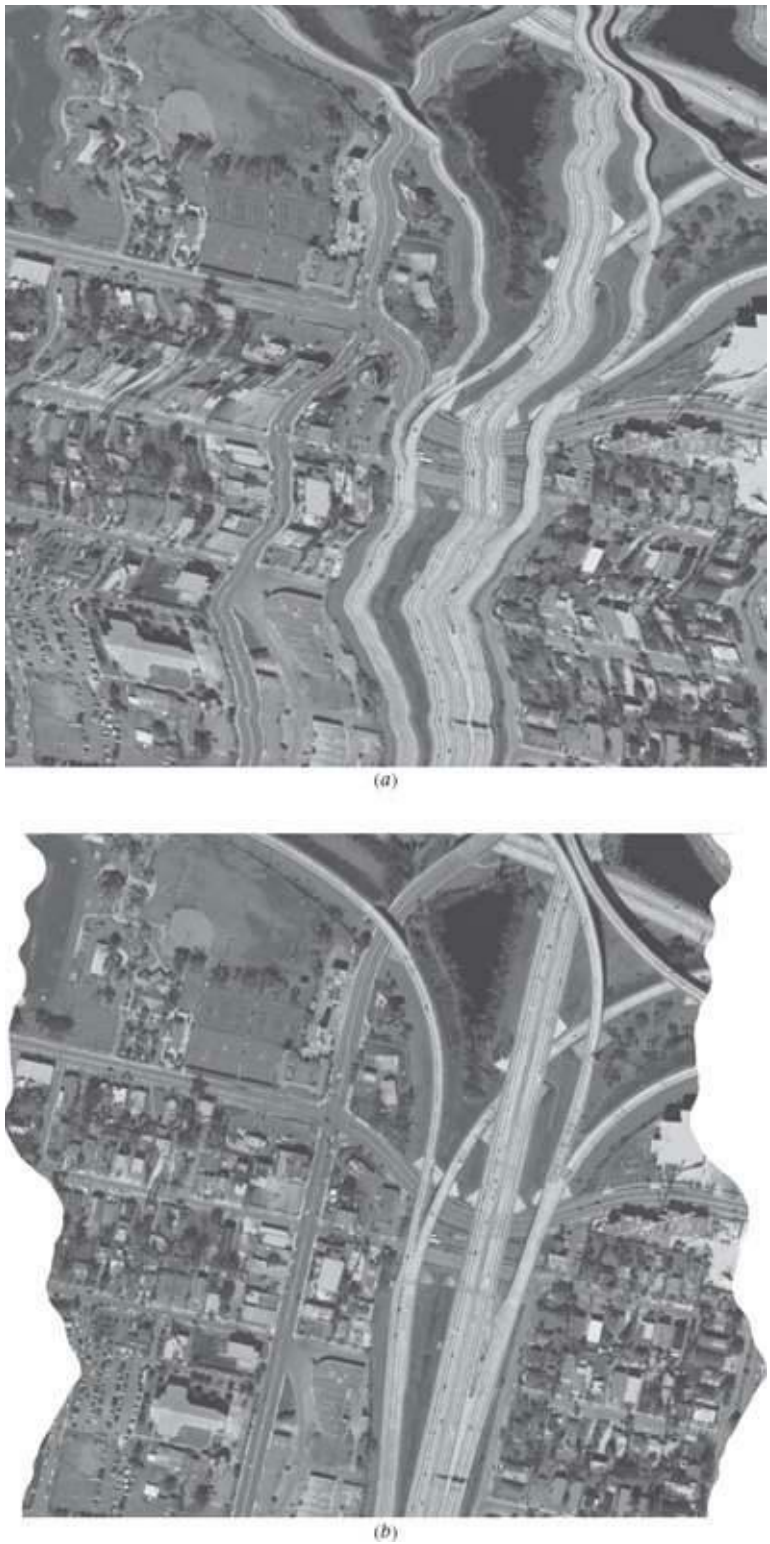


FIGURE 17-9 Raw (left) and processed (right) linear array imagery. Note that the edges of the processed imagery correspond to the tilt of the sensor during acquisition. (*Courtesy of the University of Florida*)

The first step in three-line scanner aerotriangulation is to obtain pass points between the scenes. Although pass point generation is done in Level 1 scenes to facilitate automated matching, the

coordinates of the pass points refer to the Level 0 scenes. In order to apply the collinearity equations, one must have exposure stations with multiple image observations. However, since the orientation data comes from a continuous stream, the observations of the exterior orientation parameters are continuous along the flight path and it is nearly impossible to have multiple points imaged in a single scan line. Thus, *orientation fixes* are used. Orientation fixes can be considered simulated exposure stations. They are defined at regular intervals along the flight path, and their spacing is chosen based on the quality of the GPS-INS data. The poorer the GPS-INS, the shorter the allowable interval between orientation fixes. Figure 17-10 illustrates the concept of orientation fixes along a flight path.

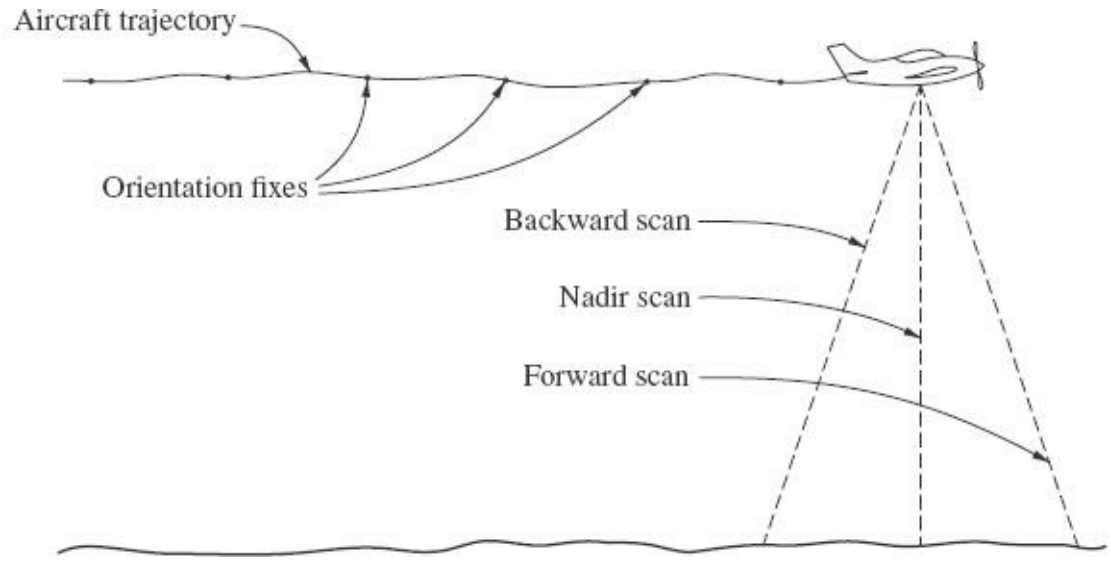


FIGURE 17-10 Orientation fixes along a flight path for a three-line linear sensor array.

Once the orientation fixes have been established, the collinearity equations for each point on each scene can be formed. The exterior orientation parameters associated with the imaging of these points must be expressed as functions of the nearest orientation fixes before and after imaging. The adjustment is similar to relative orientation in that each orientation fix for a scene is adjusted based on the weighted exterior orientation parameters of the other orientation fixes corresponding to the other scenes. Each point yields two equations for each of the three scenes. Boresight and lever arm parameters can also be introduced into the equations using methods similar to those described in Sec. 17-8. Care must be taken when selecting the distance between the orientation fixes in order to ensure that there will be enough redundancy from pass points to resolve the unknown parameters. In general, the distance between orientation fixes should not exceed the instantaneous ground distance between the nadir and backward scan lines. After the adjustment is completed, the solved orientation fixes are used to update the GPS-INS data, which can then be used to rectify Level 0 imagery.

17-11 Satellite Image Triangulation

For certain applications with low accuracy requirements, aerotriangulation from satellite images may be suitable. For example, for small-scale topographic mapping over mountainous regions, panchromatic images from a linear array sensor onboard the French *Système Pour d'Observation de la Terre* (SPOT) satellite may be used. Stereopairs of SPOT images can be acquired for a region by using the off-axis pointing capability of the satellite. Photogrammetric analysis of the resulting images can be performed through the use of modified collinearity equations.

Since the satellite is highly stable during acquisition of the image, the exterior orientation parameters can be assumed to vary in a systematic fashion. Figure 17-11 illustrates an image from a linear array sensor. In this figure, the start position (point *o*) is the projection of the center of row 0 on the ground. At this point, the satellite sensor has a particular set of exterior orientation parameters ω_o, ϕ_o, κ_o, X_{L_o}, Y_{L_o}, and Z_{L_o}. These parameters can be assumed to vary systematically as a function of the x coordinate (row in which the image appears). Various functional relationships have been tested for modeling these systematic variations, and the following have been found to consistently yield satisfactory results:

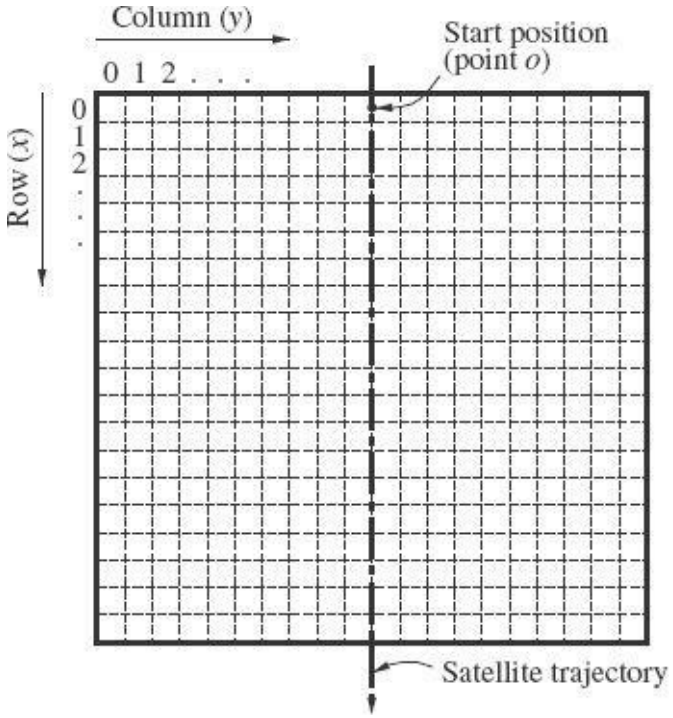


FIGURE 17-11 Illustration of linear array sensor image.

$$\begin{aligned}
\omega_x &= \omega_o + a_1 x \\
\phi_x &= \phi_o + a_2 x \\
\kappa_x &= \kappa_o + a_3 x \\
X_{L_x} &= X_{L_o} + a_4 x \\
Y_{L_x} &= Y_{L_o} + a_5 x \\
Z_{L_x} &= Z_{L_o} + a_6 x + a_7 x^2
\end{aligned} \tag{17-21}$$

In Eq. (17-21), x is the row number of some image position; ω_x, ϕ_x, κ_x, X_{L_x}, Y_{L_x}, and Z_{L_x} are the exterior orientation parameters of the sensor when row x was acquired; ω_o, ϕ_o, κ_o, X_{L_o}, Y_{L_o}, and Z_{L_o} are the exterior orientation parameters of the sensor at the start position; and a_1 through a_7 are coefficients which describe the systematic variations of the exterior orientation parameters as the image is acquired. Note that according to Eq. (17-21) the variation in Z_L is second order, whereas the other variations are linear (first order). This is due to the curved orbital path of the satellite and is based on an assumption that a local vertical coordinate system (see Sec. 5-5) is being used. Depending upon the accuracy

requirements and measurement precision, the coefficient of the second-order term a_7 may often be assumed to be equal to zero.

Given the variation of exterior orientation parameters described above, the collinearity equations which describe linear array sensor geometry for any image point *a* are

$$0 = -f\left[\frac{m_{11_x}(X_A - X_{L_x}) + m_{12x}(Y_A - Y_{L_x}) + m_{13_x}(Z_A - Z_{L_x})}{m_{31_x}(X_A - X_{L_x}) + m_{32x}(Y_A - Y_{L_x}) + m_{33_x}(Z_A - Z_{L_x})}\right] \tag{17-22}$$

$$y_a = y_o - f\left[\frac{m_{21_x}(X_A - X_{L_x}) + m_{22x}(Y_A - Y_{L_x}) + m_{23_x}(Z_A - Z_{L_x})}{m_{31_x}(X_A - X_{L_x}) + m_{32x}(Y_A - Y_{L_x}) + m_{33_x}(Z_A - Z_{L_x})}\right] \tag{17-23}$$

In Eqs. (17-22) and (17-23), y_a is the *y* coordinate (column number) of the image of point *A*; y_o is the *y* coordinate of the principal (middle) point of the row containing the image; *f* is the sensor focal length; m_{11_x} through m_{33_x} are the rotation matrix terms [see Eqs. (C-33)] for the sensor attitude when row x_a was acquired; X_{L_x}, Y_{L_x}, and Z_{L_x} are the coordinates of the sensor when row x_a was acquired; and X_A, Y_A, and Z_A are the object space coordinates of point *A*. Note that the exterior orientation terms and hence the rotation matrix terms are functions of the form of Eq. (17-21). It is also important to note that the units of the image coordinates and the focal length must be the same. For example, the first three SPOT sensor systems had focal lengths of 1082 mm and, when operating in the panchromatic mode, pixel dimensions of 0.013 mm in their focal planes.[2] Therefore, if standard row and column image coordinates (in terms of pixels) are used, the focal length is expressed as 1082 mm/0.013 mm/pixel = 83,200 pixels.

Rational polynomial coefficient (RPC) camera models (see Sec. C-10) are commonly used to describe satellite imagery. RPCs are considered a *replacement model* for the actual physical characteristics and orientation of the sensor with respect to image coordinates of ground points. They are derived from the physical model of the satellite sensor using least squares techniques, and their coefficients are delivered with the imagery. Much like the collinearity equations, RPCs are a mathematical model for transforming three-dimensional ground points to two-dimensional image coordinates. Thus, RPCs can be used in many of the same applications as the collinearity equations such as DEM generation, othorectification, and feature extraction. For example, IKONOS satellite imagery uses the ratio of two cubic polynomial functions of three-dimensional ground coordinates to describe *x* (line) and *y* (sample) coordinates of a point in the linear array sensor image as in Eq. (17-24). The image and ground coordinates of the points are normalized to avoid ill-conditioning and increase the numerical precision (see Example B-6).

$$x_a = \frac{\text{Num}_L(P_a, L_a, H_a)}{\text{Den}_L(P_a, L_a, H_a)}$$

$$y_a = \frac{\text{Num}_S(P_a, L_a, H_a)}{\text{Den}_S(P_a, L_a, H_a)} \tag{17-24}$$

In Eq. (17-24), P_a, L_a, and H_a are the normalized latitude, longitude, and height of point *a*, x_a and y_a are the normalized image coordinates of point *a*, and Num_L, Den_L, Num_S, and Den_S are cubic polynomial functions of P_a, L_a, and H_a. Both of the two rational polynomials consist of 39 coefficients (20 in the

numerator and 19 in the denominator) for a total of 78 coefficients used in the model. Note that if a point is imaged on two stereo satellite images, the three-dimensional object space coordinates can be found via least squares since there would be four equations and three unknowns, similar to space intersection via collinearity described in Sec. 11-7.

The RPC model on its own may be sufficient for some applications, however it is possible to increase the accuracy by determining bias parameters using a least squares block adjustment of stereo satellite imagery. Equation (17-25) is referred to as the adjustable RPC model, where a_0, a_1, a_2, b_0, b_1, and b_2 are affine transformation parameters that model biases in image space stemming from systematic errors in the physical orientation of the sensor.

$$x_a + a_0 + a_1 x_a + a_2 y_a = \frac{\text{Num}_L(P_a, L_a, H_a)}{\text{Den}_L(P_a, L_a, H_a)}$$

$$y_a + b_0 + b_1 x_a + b_2 y_a = \frac{\text{Num}_S(P_a, L_a, H_a)}{\text{Den}_S(P_a, L_a, H_a)} \tag{17-25}$$

The solution for the affine parameters can be found using a block adjustment of stereo satellite images with Eq. (17-25) serving as the basis for the observation equations. Depending on the geometry of the imagery and the type of sensor (e.g., IKONOS versus QuickBird) not all of the additional parameters may be statistically significant, and care should be taken not to over-parameterize the adjustment (see Sec. C-10).

17-12 Efficient Computational Strategies for Aerotriangulation

Section 17-6 presented the mathematical form for least squares adjustment of photogrammetric blocks. As presented, however, the equations are somewhat inefficient for computational purposes. Methods are available for reducing the matrix storage requirements and solution time for large blocks of photographs. One approach is briefly presented here.

The first step is to partition the full normal equations [Eq. (17-14)] so that the exterior orientation terms and the object space coordinate terms are separated, giving

$$\begin{bmatrix} \dot{N} & \bar{N} \\ \bar{N}^T & \ddot{N} \end{bmatrix} \begin{bmatrix} \dot{\Delta} \\ \ddot{\Delta} \end{bmatrix} = \begin{bmatrix} \dot{K} \\ \ddot{K} \end{bmatrix} \tag{17-26}$$

In Eq. (17-26), $\dot{N}$ is the block-diagonal submatrix from the upper left portion of N having dimensions $6m \times 6m$, where m is the number of photos in the block; $\ddot{N}$ is the block-diagonal submatrix from the lower right portion of N having dimensions $3n \times 3n$, where n is the number of object points in the block; $\bar{N}$ is the submatrix from the upper right portion of N having dimensions $6m \times 3n$ and $\bar{N}^T$ is its transpose; $\dot{\Delta}$ is the submatrix from the upper portion of Δ having dimensions of $6m \times 1$, consisting of the correction terms for the exterior orientation parameters for all photos; $\ddot{\Delta}$ is the submatrix from the lower portion of Δ having dimensions of $3n \times 1$, consisting of the correction terms for the object space coordinates for all points; $\dot{K}$ is the submatrix from the upper portion of K having dimensions of $6m \times 1$; and $\ddot{K}$ is the submatrix from the lower portion of K having dimensions of $3n \times 1$.

A block-diagonal matrix consists of nonzero submatrices along the main diagonal and zeros everywhere else. This kind of matrix has the property that its inverse is also block-diagonal, where the

submatrices are inverses of the corresponding submatrices of the original matrix. As such, the inverse of a block-diagonal matrix is much easier to compute than the inverse of a general, nonzero matrix. With this in mind, Eq. (17-26) can be rearranged to a form which can be solved more efficiently. First, Eq. (17-26) is separated into two separate matrix equations.

$$\dot{N}\dot{\Delta} + \bar{N}\ddot{\Delta} = \dot{K} \tag{17-27}$$

$$\bar{N}^T\dot{\Delta} + \ddot{N}\ddot{\Delta} = \ddot{K} \tag{17-28}$$

Equation (17-28) is then rearranged to solve for $\ddot{\Delta}$.

$$\ddot{\Delta} = \ddot{N}^{-1}(\ddot{K} - \bar{N}^T\dot{\Delta}) \tag{17-29}$$

Next the right side of Eq. (17-29) is substituted for $\ddot{\Delta}$ in Eq. (17-27).

$$\dot{N}\dot{\Delta} + \bar{N}\ddot{N}^{-1}(\ddot{K} - \bar{N}^T\dot{\Delta}) = \dot{K} \tag{17-30}$$

Rearranging Eq. (17-30) to collect the $\dot{\Delta}$ terms gives

$$(\dot{N} - \bar{N}\ddot{N}^{-1}\bar{N}^T)\dot{\Delta} = (\dot{K} - \bar{N}\ddot{N}^{-1}\ddot{K}) \tag{17-31}$$

Matrix Eq. (17-31) is referred to as the *reduced normal equations*. These equations are solved for $\dot{\Delta}$, which can then be substituted into Eq. (17-29) to compute $\ddot{\Delta}$. This approach is more efficient since the largest system of equations which must be solved has only *6m* unknowns, as opposed to *6m* + 3*n* unknowns in the full normal equations. This efficiency is made possible by the block-diagonal structure of the $\ddot{N}$ matrix.

One can also use the partitioned *N* matrix to obtain the covariance matrix. The inverse of *N* can be partitioned as shown in Eq. (17-32).

$$N^{-1} = C = \begin{bmatrix} C_1 & C_{12} \\ C_{12}^T & C_2 \end{bmatrix} \tag{17-32}$$

Using the relationship between a matrix and its inverse shown in Eq. (17-33), the matrix $C = N^{-1}$ can be formed using the definitions in Eqs. (17-34), (17-35), and (17-36).

$$CN = NC = I \tag{17-33}$$

$$C_1 = (\dot{N} - \bar{N}\ddot{N}^{-1}\bar{N}^T)^{-1} \tag{17-34}$$

$$C_{12} = -C_1\bar{N}\ddot{N}^{-1} \tag{17-35}$$

$$C_2 = \ddot{N}^{-1} + \ddot{N}^{-1}\bar{N}^T C_1 \bar{N}\ddot{N}^{-1} \tag{17-36}$$

While C_2 can be used to form the full covariance matrix for point coordinates, the computations are normally limited to determining covariance values for each point separately. This can be done by using only the portions of the matrices on the right hand side of Eq. (17-36) corresponding to a particular point *j*. The covariance matrix can then be formed using Eq. (17-16).

An additional enhancement to the solution can be made to increase computational efficiency even further. This enhancement exploits the fact that the coefficient matrix of the reduced normal equations is *sparse;* i.e., it has a large number of elements that are zero. Special computational techniques and data storage methods are available which take advantage of sparsity, reducing both computational time and data storage requirements. Details concerning these special computational techniques may be found in references listed at the end of this chapter.

Figure 17-12 shows a small block with three strips of nine photos each, having end lap and side lap equal to 60 and 30 percent, respectively. The outlines of photo coverage for only the first three photos in strips 1 and 2 are shown in the figure, and the remainder are represented as neat models (see Sec. 18-7). In Fig. 17-12, the image of a representative pass point *A* exists on photos 1-1, 1-2, 1-3, 2-1, 2-2, and 2-3. This pass point causes "connections" between each possible pair of photos from the set of six on which it is imaged. Connections for the entire block are illustrated in Fig. 17-13. This figure shows a *graph* which indicates the connections (shown as lines or arcs) caused by shared pass points over the entire block.

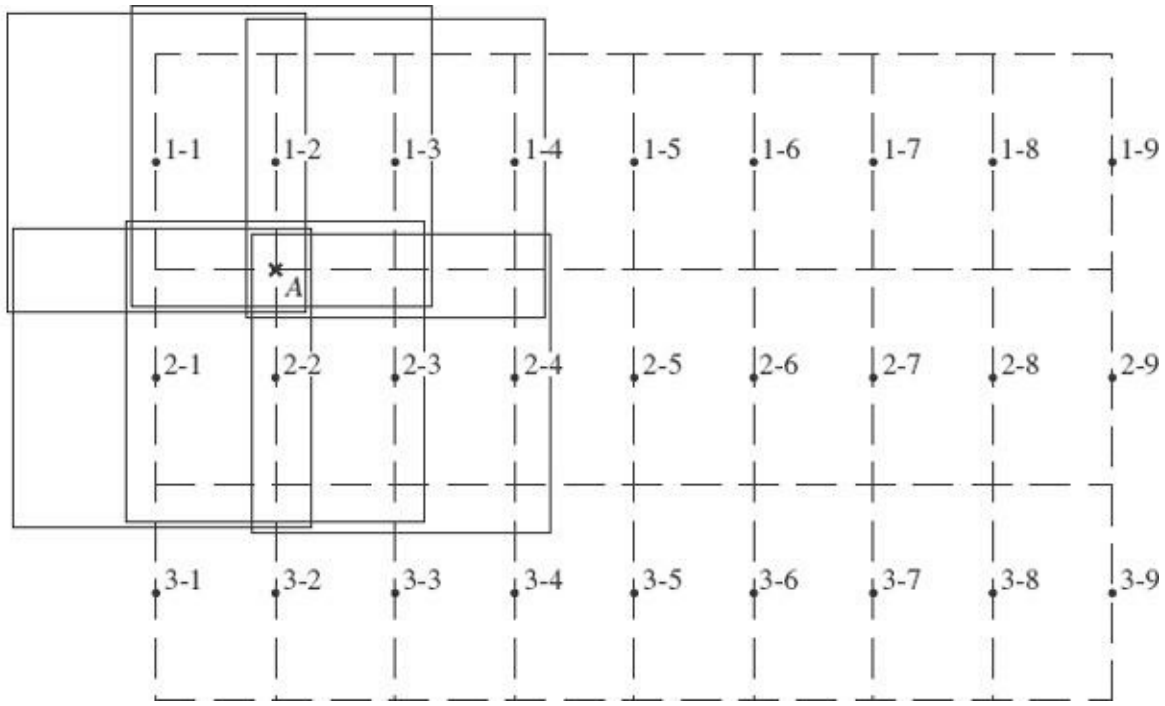


FIGURE 17-12 Configuration of a photo block having three strips of nine photos each.

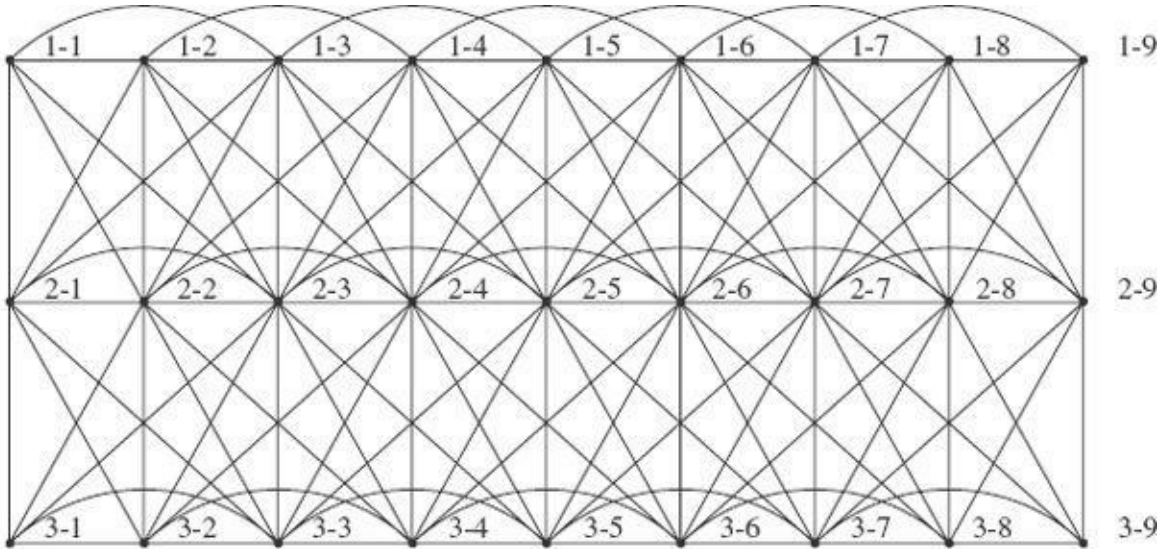


FIGURE 17-13 Graph showing connections between photos caused by shared pass points.

These connections cause nonzero submatrices to appear at corresponding locations in the reduced normal equations. The positions where these nonzero submatrices appear depend upon the order in which the photo parameters appear in the reduced normal equation matrix. Two ordering strategies, known as *down-strip* and *cross-strip*, are commonly employed. In the down-strip ordering, the photo parameters are arranged by strips, so that the nine photos from strip 1 appear first, followed by the nine photos of strip 2, and the nine photos from strip 3. With cross-strip ordering, the photo parameters are arranged so that the first photo of strip 1 appears first, followed by the first photos of strips 2 and 3; then the second photos of strips 1, 2, and 3; and so on. These two photo orders are listed in Table 17-1. As will be demonstrated, cross-strip ordering leads to a more efficient solution than down-strip ordering in this case.

Position	Down-Strip Order	Cross-Strip Order	Position	Down-Strip Order	Cross-Strip Order
1	1-1	1-1	15	2-6	3-5
2	1-2	2-1	16	2-7	1-6
3	1-3	3-1	17	2-8	2-6
4	1-4	1-2	18	2-9	3-6
5	1-5	2-2	19	3-1	1-7
6	1-6	3-2	20	3-2	2-7
7	1-7	1-3	21	3-3	3-7
8	1-8	2-3	22	3-4	1-8
9	1-9	3-3	23	3-5	2-8
10	2-1	1-4	24	3-6	3-8
11	2-2	2-4	25	3-7	1-9
12	2-3	3-4	26	3-8	2-9
13	2-4	1-5	27	3-9	3-9
14	2-5	2-5			

TABLE 17-1 Down-Strip and Cross-Strip Ordering for the Photos of Fig. 17-14

Figure 17-14 shows a schematic representation of the reduced normal equations when down-strip ordering is employed. Notice from the figure that the nonzero elements tend to cluster in a band about the main diagonal of the matrix. The width of the band from the diagonal to the farthest off-diagonal nonzero element is the *bandwidth* of the matrix. The bandwidth of the matrix shown in Fig. 17-14 is 6 × 12 = 72. With cross-strip ordering of the photos, the reduced normal equation matrix shown in Fig. 17-15 results. Here, the bandwidth is 6 × 8 = 48, which is substantially smaller than that for down-strip ordering. The narrower the bandwidth, the faster the solution and the less storage required.

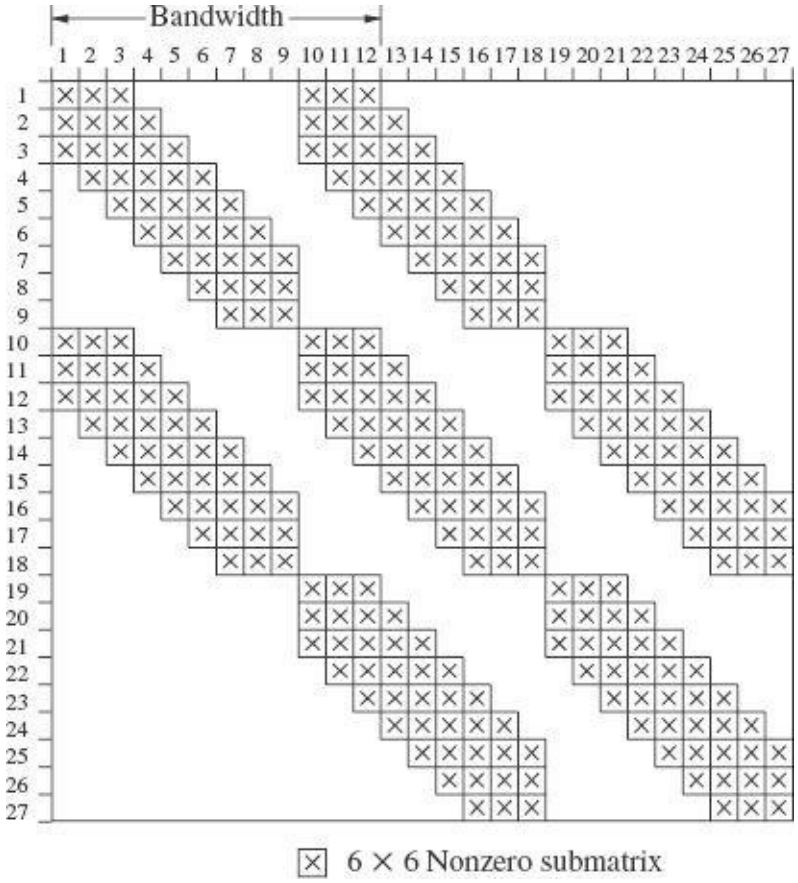


FIGURE 17-14 Structure of the reduced normal equations using down-strip ordering.

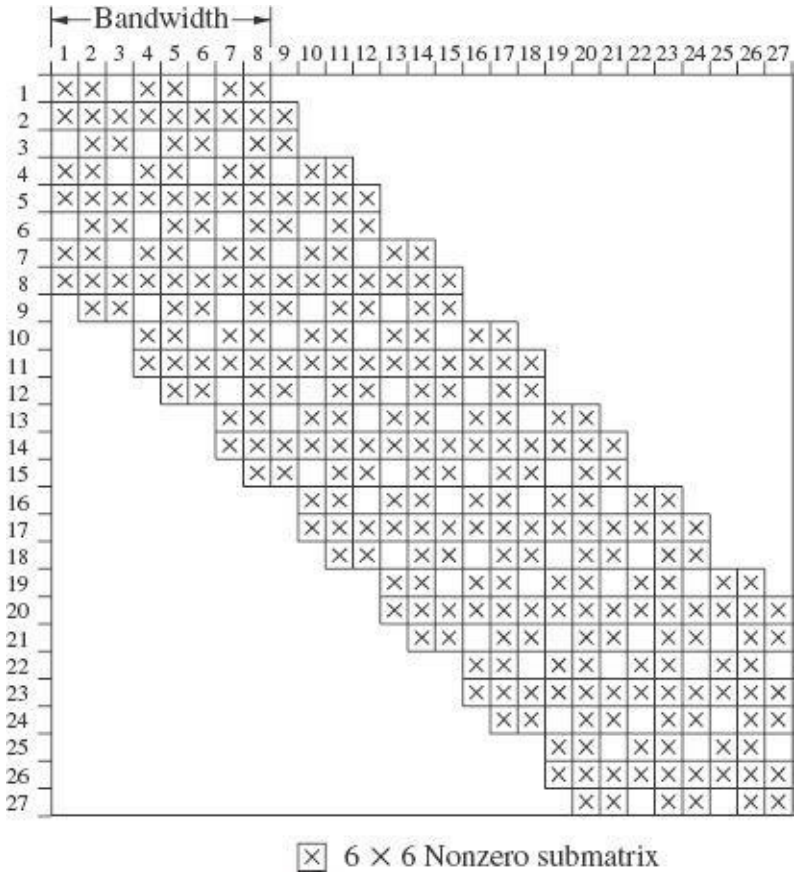


FIGURE 17-15 Structure of the reduced normal equations using cross-strip ordering.

Solution time for nonbanded reduced normal equations is proportional to the number of unknowns ($6m$) raised to the third power. For the example with 27 photos, the time is proportional to $(6 \times 27)^3 = 4.2 \times 10^6$. For banded equations, the solution time is proportional to the bandwidth squared, times the number of unknowns. For the example with down-strip number, the time is proportional to $72^2 \times (6 \times 27) = 8.4 \times 10^5$, which is 5 times faster than the nonbanded case. With cross-strip numbering, the time is proportional to $48^2 \times (6 \times 27) = 3.7 \times 10^5$, which is more than 11 times faster than the nonbanded case!

Down-strip and cross-strip ordering generally apply only to regular, rectangular photo blocks. In cases where photo blocks cover irregular areas, other more complicated approaches should be used to achieve a minimal bandwidth. Details of these other approaches can be found in references which follow.

References

Ackermann, F., and H. Schade: "Application of GPS for Aerial Triangulation," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 11, 1993, p. 1625.

American Society of Photogrammetry: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Brown, D. C.: "New Developments in Photogeodesy," *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 7, 1994, p. 877.

Curry, S., and K. Schuckman: "Practical Considerations for the Use of Airborne GPS for Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 11, 1993, p. 1611.

Duff, I. S., A. M. Erisman, and J. K. Reid: *Direct Methods for Sparse Matrices*, Oxford University Press, New York, 1990.

Ebadi, H., and M. A. Chapman: "GPS-Controlled Strip Triangulation Using Geometric Constraints of Man-Made Structures," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 4, 1998, p. 329.

El-Hakim, S. F., and H. Ziemann: "A Step-by-Step Strategy for Gross-Error Detection," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 6, 1984, p. 713.

Erio, G.: "Three-Dimensional Transformations of Independent Models," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 9, 1975, p. 1117.

Fraser, C. S., and H.T. Hanley: "Bias-Compensated RPCs for Sensor Orientation of High-Resolution Satellite Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 71, 2005, p. 909.

George, A., and J. W. H. Liu: *Computer Solution of Large Sparse Positive-Definite Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

Goad, C. C., and M. Yang: "A New Approach to Precision Airborne GPS Positioning for Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 9, 1997, p. 1067.

Grodecki, J., and G. Dial: "Block Adjustment of High-Resolution Satellite Images Described by Rational Polynomials," *Photogrammetric Engineering and Remote Sensing*, vol. 69, no. 1, 2003, p. 59.

Gruen, A., M. Cocard, and H. G. Kahle: "Photogrammetry and Kinematic GPS: Results of a High Accuracy Test," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 11, 1993, p. 1643.

Hinsken, L., S. Miller, U. Tempelmann, R. Uebbing, and S. Walker : "Triangulation of LH Systems' ADS40.

Imagery Using ORIMA GPS/IMU,"*International Archives of Photogrammetry and Remote Sensing*, vol. 34, 2001, p. 156.

Jacobsen, K.: "Experiences in GPS Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 11, 1993, p. 1651.

Kubik, K., D. Merchant, and T. Schenk: "Robust Estimation in Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 53, no. 2, 1987, p. 167.

Novak, K.: "Rectification of Digital Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 3, 1992, p. 339.

Schut, G. H.: "Development of Programs for Strip and Block Adjustment at the National Research Council of Canada," *Photogrammetric Engineering*, vol. 30, no. 2, 1964, p. 283.

Schwarz, K. P., M. A. Chapman, M. W. Cannon, and P. Gong: "An Integrated INS/GPS Approach to the Georeferencing of Remotely Sensed Data," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 11, 1993, p. 1667.

Theodossiou, E. I., and I. J. Dowman: "Heighting Accuracy of SPOT," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 12, 1990, p. 1643.

Toth, C. K., and A. Krupnik: "Concept, Implementation, and Results of an Automatic Aerotriangulation System," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 6, 1996, p. 711.

Triggs, W., P. McLauchlan, R. Hartley, and A. Fitzgibbon: "Bundle Adjustment: A Modern Synthesis," *Lecture Notes in Computer Science*, vol. 1883, 2000, p. 298.

Westin, T.: "Precision Rectification of SPOT Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 2, 1990, p. 247.

Wolf, P. R.: "Independent Model Triangulation," *Photogrammetric Engineering*, vol. 36, no. 12, 1970, p. 1262.

Problems

17-1. Discuss the advantages of aerotriangulation over field surveys.

17-2. List the three categories of aerotriangulation. Which categories are currently used?

17-3. Describe the process of automatic pass point generation. What are its advantages to manual point measurement?

17-4. Briefly describe independent model aerotriangulation by the sequential method.

17-5. Why is a three-dimensional polynomial transformation often used for transforming a sequentially constructed strip model to the ground coordinate system?

17-6. A continuous strip of three stereomodels has pass points *a* through *l* and ground control points *A* through *D*. Independent model coordinates for points and exposure stations of each model are listed below along with ground coordinates of control points. Compute the ground coordinates of the pass points and exposure stations by the sequential method of semianalytical aerotriangulation. Use the three-dimensional conformal coordinate transformation program provided (see Example 17-1).

Model Coordinates (mm)											
	Model 1-2				Model 2-3				Model 3-4		
Point	x	y	z	Point	x	y	z	Point	x	y	z
a	11.927	98.922	1.290	d	16.981	76.782	6.224	g	−3.712	78.219	2.420
b	0.408	1.579	−0.146	e	3.274	−11.530	5.904	h	10.465	−8.951	−2.352
c	−8.493	−86.188	−4.963	f	−6.140	−103.096	3.469	i	16.131	−104.252	−4.790
d	105.590	89.785	1.602	g	77.668	87.013	2.271	j	103.375	96.440	−10.590
e	93.386	1.283	−1.232	h	90.672	−0.641	1.248	k	92.321	−3.017	−7.519
f	85.528	−90.316	−6.410	i	94.747	−95.805	2.383	l	91.967	−102.867	−11.447
A	51.465	72.087	−2.051	O_2	0.000	0.000	152.673	C	84.115	61.176	4.067
B	58.489	−81.976	−5.753	O_3	72.280	14.040	155.128	D	75.773	−114.133	6.293
O_1	0.000	0.000	152.673					O_3	0.000	0.000	152.673
O_2	97.629	7.555	145.933					O_4	93.633	−6.350	155.454

Ground Control Coordinates (m)			
Point	X	Y	Z
A	23,074.937	8908.986	227.403
B	23,059.473	9371.557	233.484
C	22,467.906	8879.348	254.014
D	22,499.489	9404.362	277.061

17-7. Repeat Prob. 17-6, except using the coordinates from the table below.

Model Coordinates (Pixels)											
	Model 1-2				Model 2-3				Model 3-4		
Point	x	y	z	Point	x	y	z	Point	x	y	z
a	702.92	5859.16	77.24	d	837.97	5403.92	−44.02	g	−2915.10	4898.88	39.96
b	27.75	93.63	−11.75	e	186.33	162.58	−354.56	h	−2421.50	−322.34	171.09
c	−502.19	−5102.88	−294.96	f	−210.54	−5255.62	−825.69	i	−2480.70	−5963.22	438.62
d	6254.40	5312.51	96.48	g	4599.66	5153.87	138.78	j	6121.74	5711.33	−629.94
e	5528.50	75.50	−69.92	h	5367.81	−34.03	74.94	k	5460.72	−175.97	−443.76
f	5062.82	−5349.29	−377.13	i	5608.50	−5675.23	141.55	l	5447.22	−6087.83	−677.39
A	3040.67	4266.56	−121.22	O_2	0.00	0.00	5423.91	C	2084.22	4698.33	−821.76
B	3458.19	−4856.65	−341.43	O_3	7859.93	232.88	5296.77	D	1229.24	−5659.11	−683.42
O_1	0.00	0.00	5423.91					O_3	0.00	0.00	5423.91
O_2	5037.41	116.80	5691.88					O_4	93.78	−6.36	5517.42

Ground Control Coordinates (m)			
Point	X	Y	Z
A	6,444,921.743	798,701.808	201.790
B	6,444,892.118	799,613.868	244.895
C	6,443,722.838	798,643.722	230.886
D	6,443,785.606	799,677.620	313.381

17-8. Briefly describe how the method of independent model aerotriangulation by simultaneous transformations differs from the sequential approach.

17-9. Discuss the advantages of analytical methods of photogrammetric control extension as opposed to semianalytical methods.

17-10. Briefly describe the unknowns and measurements associated with a bundle adjustment of a block of photographs.

17-11. Describe how coordinates of the GPS antenna are related to the exposure station when airborne GPS control is used to control photography.

17-12. Describe how the boresight angular attitude parameters relate the camera and the IMU.

17-13. In what coordinate system is the lever arm offset?

17-14. Briefly discuss the problem associated with the lack of synchronization of GPS fixes with camera exposures in airborne GPS control.

17-15. What is the purpose of cross strips at the ends of photo blocks when airborne GPS is used to control photography?

17-16. Compute initial approximations for the angular exterior orientation parameters using the results from a sequentially constructed strip model with results in the following table using the methods in Example 17-2.

	ω	φ	κ
Relative orientation of photo 1 to photo 2	1.3221°	−1.3549°	1.2244°
Relative orientation of photo 2 to photo 3	0.0085°	1.7933°	1.6357°
Relative orientation of photo 3 to photo 4	0.9746°	−1.4990°	1.0065°
Orientation of ground to strip	2.2396°	−1.2120°	−3.1492°

17-17. Briefly explain how a line perspective image differs from a point perspective image.

17-18. Briefly discuss the characteristic of the $\ddot{N}$ submatrix that makes the method of reduced normal equations more efficient for a bundle adjustment than solving the full normal equations.

17-19. Discuss the difference between down-strip and cross-strip numbering as they apply to the bandwidth of the reduced normal equations of a bundle adjustment.

---

[2] SPOT 1, 2, and 3 sensors could be operated in either a panchromatic or multispectral mode. In panchromatic mode, pixel dimensions were 0.013 mm, and ground resolution was 10 m. In multispectral mode, pixel dimensions were 0.026 mm, and ground resolution was 20 m.

CHAPTER 18
Project Planning

18-1 Introduction

Successful execution of any photogrammetric project requires that thorough planning be done prior to proceeding with the work. Planning, more than any other area of photogrammetric practice, must be performed by knowledgeable and experienced persons who are familiar with all aspects of the subject.

The first and most important decision to be made in the planning process concerns the selection of the products that will be prepared. In addition to selecting the products, their scales and accuracies must be fixed. These decisions can only be made if the planner thoroughly understands what the client's needs are, so that the best overall products can be developed to meet those needs. The client will also naturally be concerned with the anticipated costs of the items as well as the proposed schedule for their delivery. Therefore, successful planning will probably require several meetings with the client prior to commencing the work, and depending upon the nature and magnitude of the project, continued meetings may be needed as production progresses.

A variety of products may be developed in a given photogrammetric project, including prints of aerial photos, photo indexes, photomaps, mosaics, orthophotos, planimetric and topographic maps, digital maps for GIS databases and other purposes, cross sections, digital elevation models, cadastral maps, and others. In addition to the wide variation in products that could be developed for a given project, there are normally other major considerations that will have definite bearing on procedures, costs, and scheduling. These include the location of the project; its size, shape, topography, and vegetation cover; the availability of existing ground control; etc. Thus, every project presents unique problems to be considered in the planning stages.

Assuming that the photogrammetric products, and their scales and accuracies, have been agreed upon with the client, the balance of the work of project planning can generally be summarized in the following categories:

1. Planning the aerial photography[1]
2. Planning the ground control
3. Selecting instruments and procedures necessary to achieve the desired results
4. Estimating costs and delivery schedules

When planning has been completed for these categories, the photogrammetrist will normally prepare a detailed proposal which outlines plans, specifications, an estimate of costs, and delivery schedules for the project. The proposal often forms the basis of an agreement or contract for the performance of the work.

Of the above four categories, this chapter concentrates primarily on planning of the aerial photography. Planning of the ground control has been discussed in detail in Chap. 16, and instrument and procedure selection has been discussed in earlier chapters where the various photogrammetric products and instruments for producing them have been described. A brief discussion of the subjects of cost estimating and scheduling is given in Sec. 18-12.

18-2 Importance of Flight Planning

Because the ultimate success of any photogrammetric project probably depends more upon good-quality photography than on any other aspect, planning the aerial photography, also called *flight planning*, is of major concern. If the photography is to satisfactorily serve its intended purposes, the photographic mission must be carefully planned and faithfully executed according to the "flight plan." A flight plan generally consists of two items: a *flight map*, which shows where the photos are to be taken; and *specifications*, which outline how to take them, including specific requirements such as camera and film requirements, scale, flying height, end lap, side lap, and tilt and crab tolerances. A flight plan which gives optimum specifications for a project can be prepared only after careful consideration of all the many variables which influence aerial photography.

An aerial photographic mission is an expensive operation involving two or more crewpersons and high-priced aircraft and equipment. In addition, periods of time that are acceptable for aerial photography are quite limited in many areas by weather and ground cover conditions, which are related to seasons of the year. Failure to obtain satisfactory photography on a flight mission not only necessitates costly reflights, but also in all probability will cause long and expensive delays on the project for which the photos were ordered. For these reasons flight planning is one of the most important operations in the overall photogrammetric project.

18-3 Photographic End Lap and Side Lap

Before discussing the many aspects which enter into consideration in planning an aerial photographic mission, it will be helpful to redefine the terms *end lap* and *side lap*. As discussed in Sec. 1-4, vertical aerial photographic coverage of an area is normally taken as a series of overlapping flight strips. As illustrated in Fig. 18-1, *end lap* is the overlapping of successive photos along a flight strip. Figure 18-2 illustrates *side lap*, or the overlap of adjacent flight strips.

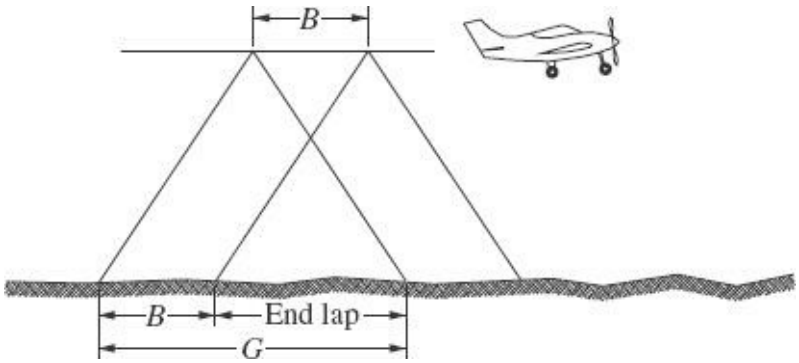


FIGURE 18-1 End lap, the overlapping of successive photos along a flight strip.

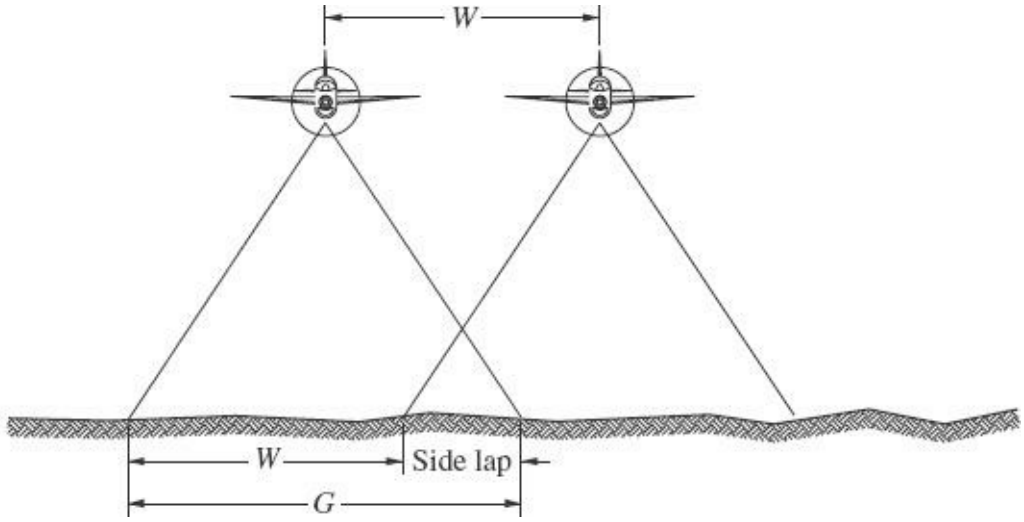


FIGURE 18-2 Side lap, the overlapping of adjacent flight strips.

In Fig. 18-1, G represents the dimension of the square of ground covered by a single vertical photograph (assuming level ground and a square camera focal-plane format), and B is the air base or distance between exposure stations of a stereopair. The amount of end lap of a stereopair is commonly given in percent. Expressed in terms of G and B, it is

$$\text{PE} = \frac{G - B}{G} \times 100 \tag{18-1}$$

In Eq. (18-1), PE is percent end lap. If stereoscopic coverage of an area is required, the absolute minimum end lap is 50 percent. However, to prevent gaps from occurring in the stereoscopic coverage due to crab, tilt, flying height variations, and terrain variations, end laps greater than 50 percent are used. Also, if the photos are to be used for photogrammetric control extension, images of some points must appear on three successive photographs—a condition requiring greater than 50 percent end lap. For these reasons aerial photography for mapping purposes is normally taken with about 60 percent end lap.

Crab, as explained in Sec. 3-6, exists when the edges of the photos in the x direction are not parallel with the direction of flight. It causes a reduction in stereoscopic coverage, as was indicated in Fig. 3-7. Figures 18-3 through 18-5 illustrate reductions in end lap causing loss of stereoscopic coverage due to tilt, flying height variation, and relief variations, respectively.

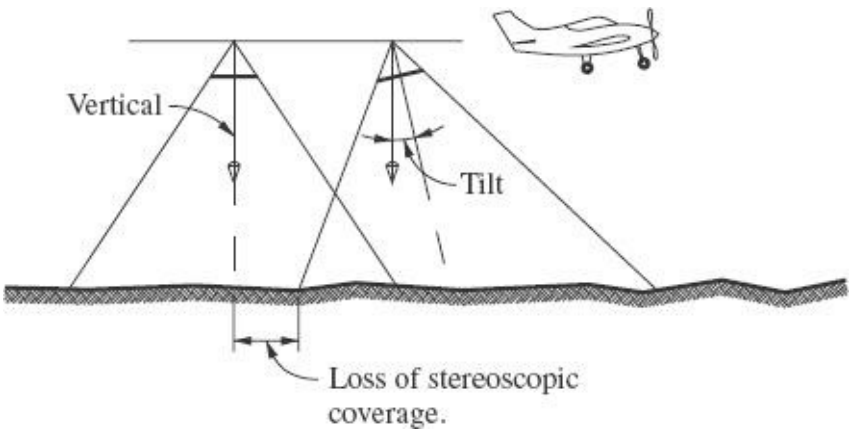


FIGURE 18-3 Failure to achieve stereoscopic coverage due to tilt.

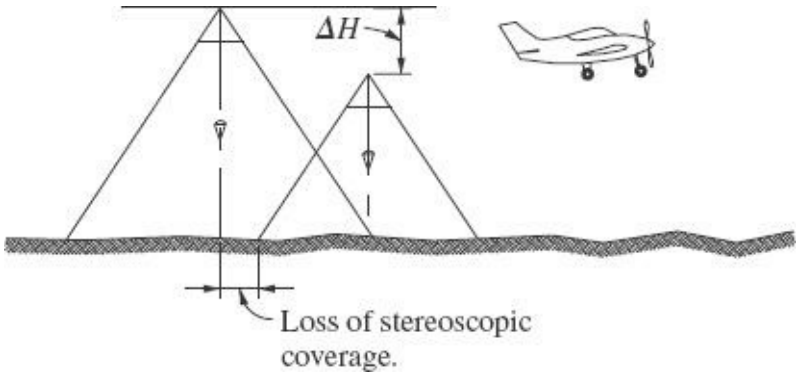


FIGURE 18-4 Failure to achieve stereoscopic coverage due to flying height variations.

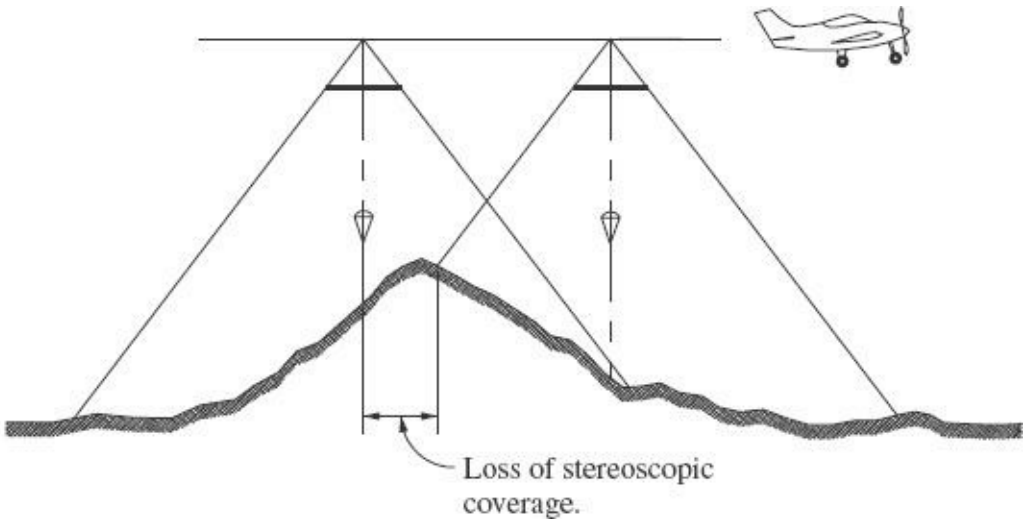


FIGURE 18-5 Failure to achieve stereoscopic coverage due to terrain variations.

Side lap is required in aerial photography to prevent gaps from occurring between flight strips as a result of drift, crab, tilt, flying height variation, and terrain variations. *Drift* is the term applied to a failure of the pilot to fly along planned flight lines. It is often caused by strong winds, but can also result from a lack of definite features and objects shown on the flight map which can also be identified from the air to guide the pilot during photography. Excessive drift is the most common cause for gaps in photo coverage; when this occurs, reflights are necessary.

In Fig. 18-2, again G represents the dimension of the square of ground coverage of a single vertical photograph, and W is the spacing between adjacent flight lines. An expression for PS (percent side lap) in terms of G and W

$$\text{PS} = \frac{G - W}{G} \times 100 \tag{18-2}$$

Mapping photography is normally taken with a side lap of about 30 percent. Besides helping to prevent gaps in coverage, another advantage realized from using this large a percentage is the elimination of the need to use the extreme edges of the photography, where the imagery is of poorer quality. Photography for orthophoto or mosaic work is sometimes taken with greater than 30 percent side lap since this reduces the sizes of the central portions of the photographs that must be used, thereby lessening distortions of images due to tilt and relief. In certain cases, such as for very precise

photogrammetric control extension, the aerial photos may be taken with 60 percent side lap as well as 60 percent end lap to increase the redundancy in the bundle adjustment (see Chap. 17).

Example 18-1
The air base of a stereopair of vertical photos is 1400 m, and flying height above average ground is 2440 m. The camera has a focal length of 152.4 mm and a 23-cm format. What is the percent end lap?

Solution

1. $$\text{Average photo scale} = \frac{f}{H'_{\text{avg}}} = \frac{152.4\,\text{mm}}{2440\,\text{m} \times 1000\,\text{mm/m}}$$
$$= 1:16{,}000 \,(1\,\text{in} = 1330\,\text{ft})$$

2. Ground coverage dimension
$$G = \frac{23\,\text{cm} \times 16{,}000}{100\,\text{cm/m}} = 3680\,\text{m}\,(12{,}100\,\text{ft})$$

3. Percent end lap, by Eq. (18-1), is
$$\text{PE} = \frac{3680\,\text{m} - 1400\,\text{m}}{3680\,\text{m}} \times 100 = 62\% \quad \blacktriangle$$

Example 18-2
In Example 18-1, assume that the spacing between adjacent flight strips is 2500 m. What is the percent side lap?

Solution By Eq. (18-2),
$$\text{PS} = \frac{3680\,\text{m} - 2500\,\text{m}}{3680\,\text{m}} \times 100 = 32\% \quad \blacktriangle$$

18-4 Purpose of the Photography

In planning aerial photographic missions, the first and foremost consideration is the purpose for which the photography is being taken. Only with the purpose defined can optimum equipment and procedures be selected. In general, aerial photographs are desired which have either good *metric* qualities or high *pictorial* qualities. Photos having good metric qualities are needed for topographic mapping or other purposes where precise quantitative photogrammetric measurements are required. High pictorial qualities are required for qualitative analysis, such as for photographic interpretation or for constructing orthophotos, photomaps, and aerial mosaics.

Photographs of good metric quality are obtained by using calibrated cameras and films having fine-grained, high-resolution emulsions or digital sensors with a high pixel count. For topographic mapping and other quantitative operations, photography is preferably taken with a wide- or super-wide-angle (short-focal-length) camera so that a large *base-height* ratio (B/H') is obtained. The B/H' ratio, as described in Sec. 7-8, is the ratio of the air base of a pair of overlapping photographs to the

average flying height above ground. The larger the B/H' ratio, the greater the intersection angles (parallactic angles) between intersecting light rays to common points. In Figs. 18-6*a* and *b*, for example, the air bases are equal, but the focal length and flying height in Fig. 18-6*a* are one-half those in Fig. 18-6*b*. The photographic scales are therefore equal for both cases, but the B/H' ratio of Fig. 18-6*a* is double that of Fig. 18-6*b*, and parallactic angle ϕ_1 to point A in Fig. 18-6*a* is nearly double the corresponding angle ϕ_2 in Fig. 18-6*b*.

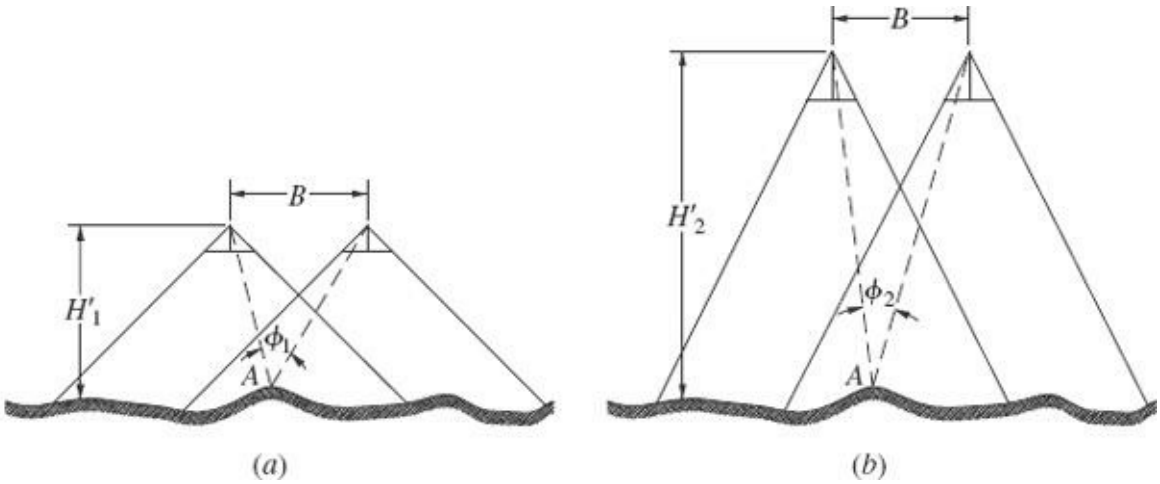


FIGURE 18-6 Parallactic angles increase with increasing B/H' ratios.

It can be shown that the errors in computed positions and elevations of points in a stereopair increase with increasing flying heights, and decrease with increasing x parallax. Large B/H' ratios denote low flying heights and large x parallaxes, conditions favorable to higher accuracy. The photos of Fig. 18-6*a* are therefore superior to those of Fig. 18-6*b* for mapping or quantitative analyses.

Photography of high pictorial quality does not necessarily require a calibrated camera, but the camera must have a good-quality lens. In many cases, films having fast, large-grained emulsions or digital cameras with high ISO speeds produce desirable effects. For some photo interpretation work, normal color photography is useful. For other special applications, black-and-white infrared or color infrared photos are desirable. Special effects can also be obtained by using filters in combination with various types of films. Timber types, for example, can be delineated quite effectively by using a red filter in combination with black-and-white infrared film. Similarly, digital image processing techniques can be applied to enhance photos for many different interpretive applications.

For mosaic work, relief displacements, tilt displacements, and scale variations produce objectionable degradations of pictorial quality. These may be minimized, however, by increasing flying height, thereby decreasing the B/H' ratio. Increased flying height, of course, reduces photo scale; but compensation for this can be achieved by using a longer-focal-length camera. The photo of Fig. 18-7*a* was exposed at one-half the flying height of the photo of Fig. 18-7*b*. The scales of the two photos are equal, however, because focal length f_2 of Fig. 18-7*b* is double f_1 of Fig. 18-7*a*. The photo o f Fig. 18-7*b* is more desirable for mosaic construction because its scale variations and image distortions due to relief, tilt, and flying height variations are much less than those of the photo of Fig. 18-7*a*. On Fig. 18-7*a*, for example, relief displacement d_1 is more than double the corresponding relief displacement d_2 of Fig. 18-7*b*.

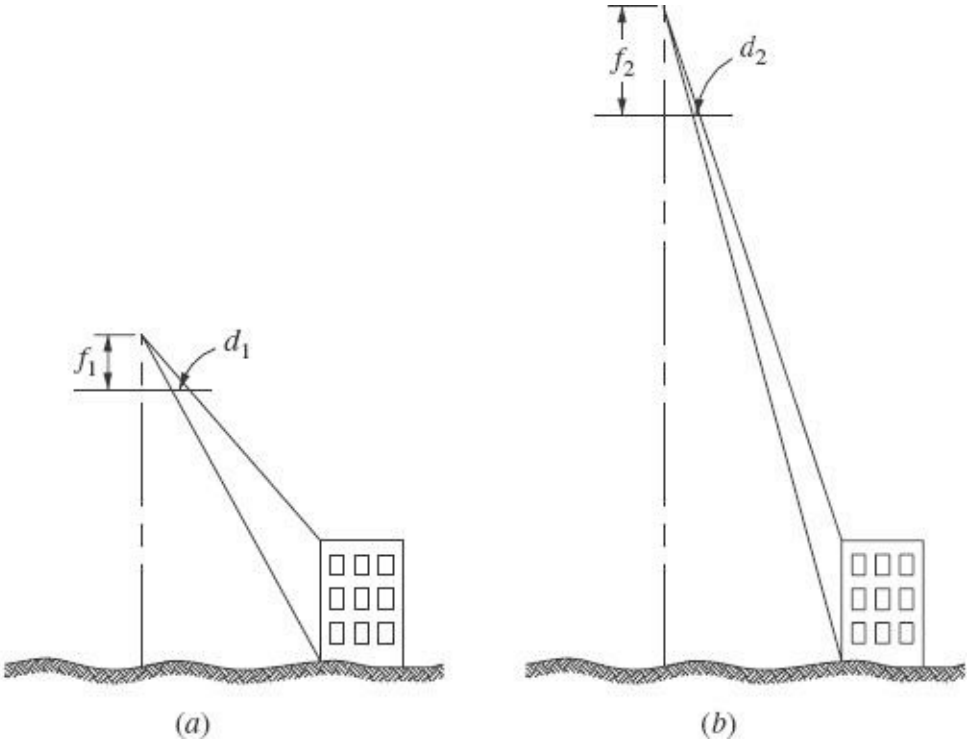


FIGURE 18-7 Reduction in relief displacement is achieved by increasing flying height.

18-5 Photo Scale

Average photographic scale is one of the most important variables that must be selected in planning aerial photography. It is generally fixed within certain limits by specific project requirements. For topographic mapping, photo scale is usually dictated by the map's required scale and/or horizontal and vertical accuracy. On the other hand, in photo interpretation applications, or for orthophoto or mosaic preparation, ensuring that the smallest ground objects of importance can be resolved on the photos may be the governing consideration in selecting the photo scale.

Example 18-3

A particular project involves the use of aerial photography to study the centerline markings of highways. The actual width of painted centerlines on the highways is 100 mm (4 in). A high-resolution film (80 line pairs per millimeter) will be used. What minimum photo scale is required?

Solution With a resolution of 80 line pairs per millimeter (see Sec. 3-14), the smallest objects that could reasonably be resolved with that film would be 1/80 mm = 0.0125 mm in size. Thus, the minimum scale required would be roughly

$$S_{\min} = \frac{0.0125\,\text{mm}}{100\,\text{mm}} = 1/8000 \text{ or } 1:8000 (1\,\text{in} = 670\,\text{ft}) \quad \blacktriangle$$

In topographic mapping, the enlargement ratio from photo scale to the scale of the plotted map must be considered. With older direct optical projection stereoplotters, where maps were drawn in real time as the operator viewed and traced the stereomodel, the enlargement ratio was fixed within narrow limits, with 5 being most common. Today these types of instruments are seldom used in practice. Mechanical projection stereoplotters (see Chap. 12) have enlargement ratio capabilities which range from less than 1 up to 10 or more, depending upon the particular instrument. However, the number of these instruments in use today is also declining, and those being used are seldom employed for direct

map compilation anymore. Rather they are equipped with digitizers, and topographic information is digitized and stored in files. From these digital files, maps are plotted using computers. Today, analytical plotters and softcopy systems account for the majority of topographic mapping, and for topographic mapping both of these systems are normally used to produce digital files.

In using computers to plot maps from digital files, virtually any enlargement ratio is possible. It is important to note, however, that the digitized information contains errors, and these errors are magnified by whatever enlargement ratio is used. Thus to ensure the integrity and accuracy of a map compiled from digital data, the enlargement ratios that are actually utilized must be held within reasonable limits. To ensure that their plotted maps meet required accuracy standards, many organizations will not enlarge more than five to seven times over photo scale. Higher enlargement ratios are sometimes used but this should be done with caution, and generally only when experience with previous projects has shown through field checks that satisfactory accuracies have been obtained.

Example 18-4
A map must be compiled at a scale of 1:6000. If an enlargement ratio of 5 will be used in producing this map, what is the required photo scale?

Solution Photo scale is one-fifth as large as map scale. Therefore,

$$S_{\text{photo}} = \frac{1}{6000} \times \frac{1}{5} = 1:30{,}000 (1\text{ in} = 2500\text{ ft}) \quad \blacktriangle$$

Selection of optimum map scale depends upon the purpose of the map. It should be carefully planned, because compilation at a larger scale than necessary is uneconomical, and compilation at too small a scale reduces the usefulness of the map or may even render it unsatisfactory. The horizontal accuracy (accuracy to which planimetric positions of points can be measured from a map) depends directly upon the map's scale. Assume, for example, that map positions of planimetric features can be plotted correctly to within 1/30 in, a condition necessary to meet National Map Accuracy Standards (see Sec. 16-1) for large-scale maps.[2] Now if a particular cadastral map requires that points be accurate to within ± 2.0 ft, required map scale is 1 in = 60 ft (1:720). Then for that situation, if an enlargement ratio of 5 is employed, photo scale is fixed at 5 × 60 = 300 ft/in (1:3600). In another example, if points only need to be accurate to within ± 20 ft on a topographic map for preliminary planning, and again assuming the same accuracy of 1/30 in, then a scale of 1 in = 600 ft (1:7200) is all that would be required, and minimum photo scale would be 5 × 7200, or 1:36,000.

As previously noted, vertical mapping accuracy is also an important factor to be considered in planning aerial photography. In past photogrammetric practice, contours which portrayed topographic relief were compiled directly from stereomodels, and thus the guidelines and standards for specifying vertical accuracy in topographic mapping were commonly given in terms of contour interval. As noted earlier in this text, however, for modern photogrammetry projects digital elevation models (DEMs) are now generally compiled rather than contours. From the DEMs, triangulated irregular network (TIN) models are constructed using computers, and then contours, cross sections, and profiles can be generated automatically from the TIN models. But even though different procedures are now employed to compile topographic information, contours are still the end product that is often used in representing topographic relief. Thus, guidelines and standards that are based on contour interval are still appropriate for quantifying vertical mapping accuracy. And as will be discussed later in this section, as vertical mapping accuracy requirements increase (contour interval decreases), flying height must decrease and photo scale must increase.

As with planimetric accuracy, the contour interval to be selected for a particular mapping project

depends upon the intended use of the maps. Assume, for example, that elevations can be interpolated correctly from a map to within one-half the contour interval, a condition required for meeting National Map Accuracy Standards. If elevations must be interpolated to within ±0.5 ft on a highway design map, then a 1-ft contour interval is necessary. If elevations must be interpolated to ±10 ft on a map prepared for studying the volume of water impounded in the reservoir of a large dam, then a 20-ft contour interval is all that is required.

The recommended contour interval depends on not only the use to be made of the map but also the type of terrain. If the map is being prepared for planning a sewer system for a city such as Las Vegas, Nevada, which lies on relatively flat terrain, perhaps a 1-ft contour interval would be required. On the other hand, if a topographic map of San Francisco is being prepared for the same purpose, because of the large range of relief in that city, perhaps a 5- or 10-ft contour interval would be used.

In planning a topographic mapping project, contour interval and map scale must be selected so that they are compatible. As map scale decreases, contour interval must increase; otherwise, the contours would become too congested on the map. In large-scale mapping of average types of terrain, the scale and contour interval relationships shown in Table 18-1 generally provide satisfactory compatibility.

English System		Metric System	
Map Scale	**Contour Interval**	**Map Scale**	**Contour Interval**
1 in/50 ft	1 ft	1:500	0.5 m
1 in/100 ft	2 ft	1:1000	1 m
1 in/200 ft	5 ft	1:2000	2 m
1 in/500 ft	10 ft	1:5000	5 m
1 in/1000 ft	20 ft	1:10,000	10 m

TABLE 18-1 Compatible Map Scales and Contour Intervals for Average Terrain

Relative accuracy capabilities in photogrammetric mapping, whether planimetric or vertical, depend upon many variables, but the most important is flying height above ground. Others include the quality of the stereoplotting instrument that is used and the experience and ability of its operator, the camera and its calibration, the quality of the photography, the density and accuracy of the ground control, and the nature of the terrain and its ground cover. A rule of thumb for quantifying vertical accuracy capability, based on contour interval, employs a term called the *C factor*. The *C* factor is the ratio of the flying height above ground of the photography (H') to the contour interval that can be reliably plotted using that photography, or in equation form

$$\text{C factor} = \frac{H'}{\text{CI}} \tag{18-3}$$

The units of H' and CI (contour interval) of Eq. (18-3) are the same. The *C* factors that are employed by photogrammetric mapping organizations are based upon their experiences, and this experience will include field checks of map accuracies achieved on a variety of previous projects. To ensure that their maps meet required accuracy standards, many organizations use a *C* factor of from about 1200 to 1500. Other organizations may push the value somewhat higher, but this must be done with extreme

caution.

Example 18-5
A topographic map having a scale of 200 ft/in with 5-ft contour interval is to be compiled from contact-printed diapositives using a stereoplotter having a nominal 6-in (152-mm) principal distance. Determine the required flying height for the photography if the maximum values to be employed for the *C* factor and enlargement ratio are 1500 and 5, respectively.

Solution

1. Considering contour interval and *C* factor:

$$H' = 1500(5\text{ ft}) = 7500\text{ ft}$$

2. Considering map scale and enlargement ratio:

$$\text{Photo scale} = \frac{1\text{ in}}{200\text{ ft}} \times \frac{1}{5} = \frac{1\text{ in}}{1000\text{ ft}} = \frac{f}{H'}$$

$$\text{Thus,} \qquad H' = 6\text{ in} \times 1000\text{ ft/in} = 6000\text{ ft}$$

In this instance, the lower of the two flying heights (6000 ft) must be selected, and therefore the enlargement ratio from photo scale to map scale governs over contour interval. ▲

In some topographic mapping projects, particularly where there is little relief, it is impractical to show elevations using contours because few, if any, may result. In these situations, a grid of spot elevations can be read throughout the area to depict the relief. A rule of thumb relating accuracy of spot elevations and flying height is that the ratio of flying height above ground to the accuracy with which spot elevations can be reliably read is approximately 5000. Thus, if spot elevations are needed to an accuracy of ±1/2 m, then the flying height above ground necessary to achieve those results would be in the range of about 1/2 (5000), or roughly 2500 m. Again, in addition to flying height, the actual accuracies that can be achieved also relate to the stereoplotting instrument, the operator's ability, the quality of the aerial photography, the density and accuracy of the ground control, and other factors.

18-6 Flying Height

Once the camera focal length and required average photo scale have been selected, required flying height above average ground is automatically fixed in accordance with scale [see Eq. (6-3)].

Example 18-6
Aerial photography having an average scale of 1:6000 is required to be taken with a 152.4-mm-focal-length camera over terrain whose average elevation is 425 m above mean sea level. What is required flying height above mean sea level?

Solution By Eq. (6-3)

$$S = \frac{f}{H - h_{avg}}$$

Thus $$\frac{1}{6000} = \frac{152.4\,\text{mm}}{H - 425\,\text{m}}\left(\frac{1\,\text{m}}{1000\,\text{mm}}\right)$$

$$H = 6000(0.1524\,\text{m}) + 425\,\text{m} = 1340\,\text{m above MSL} \quad \blacktriangle$$

Flying heights above average ground may vary from a hundred meters or so in the case of large-scale helicopter photography, to several hundred kilometers if satellites are used to carry the camera. Flying heights used in taking photos for topographic mapping normally vary between about 500 and 10,000 m. If one portion of the project area lies at a substantially higher or lower elevation than another part, two different flying heights may be necessary to maintain uniform photo scale.

Ground coverage per photo for high-altitude photography is greater than that for low-altitude photography. Fewer high-altitude photos are therefore required to cover a given area. Very high-altitude coverage is more expensive to obtain than low-altitude photography because of the special equipment required. Some of the problems encountered at high flying heights are the decreasing available oxygen, decreasing pressure, and extreme cold. When flying heights exceed about 3000 m, an oxygen supply system is necessary for the flight crew. At altitudes above 10,000 m, pure oxygen under pressure is required. Also, the cabin must be pressurized, and heaters are required to protect the crew against the cold. Most aerial photography is taken by using single- or twin-engine aircraft. Supercharged single-engine aircraft can reach altitudes of about 6 km, and supercharged twin-engine aircraft are capable of approaching 10 km. Higher altitudes require turbocharged or jet aircraft.

During photography the pilot maintains proper flying height by means of an altimeter or GPS receiver. Since altimeters give elevations above mean sea level, the proper reading is the sum of average ground elevation and required flying height above ground necessary to achieve proper photo scale. Altimeters are barometric instruments, and consequently their readings are affected by varying atmospheric pressure. They must be checked daily and adjusted to base airport air pressure. GPS receivers, while essentially unaffected by barometric pressure, give elevations relative to the ellipsoid (see Sec. 16-5); therefore a geoid model is required to relate these elevations to mean sea level.

18-7 Ground Coverage

Once average photographic scale and camera format dimensions have been selected, the ground surface area covered by a single photograph may be readily calculated. In addition, if end lap and side lap are known, the ground area covered by the stereoscopic *neat model* can be determined. The neat model, as illustrated in Fig. 18-8, is the stereoscopic area between adjacent principal points and extending out sideways in both directions to the middle of the side lap. The neat model has a width of B and a breadth of W. Its coverage is important since it represents the approximate mapping area of each stereopair.

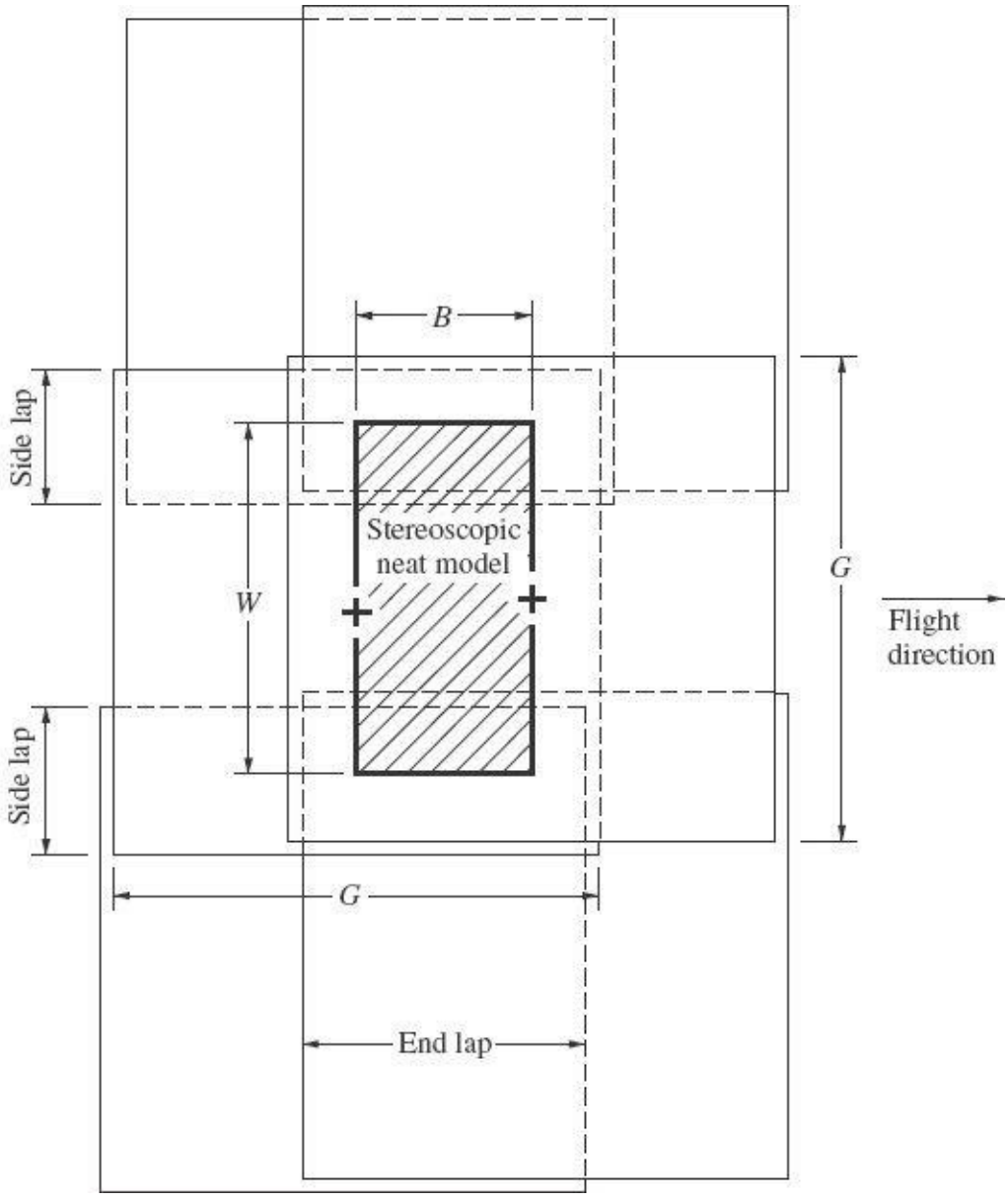


FIGURE 18-8 The area covered by a stereoscopic neat model.

Example 18-7

Aerial photography is to be taken from a flying height of 6000 ft above average ground with a camera having a 6-in (152.4-mm) focal length and a 9-in (23-cm) format. End lap will be 60 percent, and side lap will be 30 percent. What is the ground area covered by a single photograph and by the stereoscopic neat model?

Solution

1. By Eq. (6-1),

$$S = \frac{6 \text{ in}}{6000 \text{ ft}} = \frac{1 \text{ in}}{1000 \text{ ft}} \text{ or } 1:12{,}000$$

2. The dimension G of the square ground area covered by a single photo is

$$G = 1000 \text{ ft/in} \times 9 \text{ in} = 9000 \text{ ft}$$

3. The area in acres covered on the ground by a single photo is

$$A = \frac{(9000\ \text{ft})^2}{43{,}560\ \text{ft}^2/\text{acre}} = 1900\ \text{acres}$$

4. At 60 percent end lap, B is $0.4G$; and at 30 percent side lap, W is $0.7G$. Therefore the dimensions of the rectangular stereoscopic neat model are

$$B = 0.4(9000\ \text{ft}) = 3600\ \text{ft}$$
$$W = 0.7(9000\ \text{ft}) = 6300\ \text{ft}$$

And the area of the neat model is

$$\frac{3600\ \text{ft} \times 6300\ \text{ft}}{43{,}560\ \text{ft}^2/\text{acre}} = 520\ \text{acres} \quad \blacktriangle$$

While neat models guide the planning process for projects intended for stereoplotter mapping, other applications may call for different ground coverage parameters. For example, if orthophotos are the intended final product, it is beneficial to plan flights with equal end lap and side lap. The advantage of increasing the side lap is that more ground is covered near the center of photos. This reduces the amount of relief displacement and occlusion by buildings, and is therefore more useful in urban areas than rural.

18-8 Weather Conditions

The weather, which in most locations is uncertain for any given day, is a very important consideration in aerial photography. In most cases, an ideal day for aerial photography is one that is free from clouds, although if the sky is less than 10 percent cloud-covered, the day may be considered satisfactory. If clouds of greater than 10 percent coverage are present but are so high that they are above the planned flying height, this may still be objectionable since large cloud shadows will be cast on the ground, obscuring features. The number of satisfactory cloudless days varies with time of year and locality. There are certain situations in which overcast weather can be favorable for aerial photography. This is true, for example, when large-scale photos are being taken for topographic mapping over built-up areas, forests, steep canyons, or other features which would cast troublesome shadows on clear, sunny days.

A particular day can be cloudless and still be unsuitable for aerial photography due to atmospheric haze, smog, dust, smoke, high winds, or air turbulence. Atmospheric haze scatters almost entirely in the blue portion of the spectrum, and it can therefore be effectively eliminated from the photographs by using a yellow filter in front of the camera lens. Smog, dust, and smoke scatter throughout the entire spectrum and cannot be filtered out satisfactorily. Best days for photographing over industrial areas which are susceptible to smog, dust, and smoke occur after heavy rains or during moving cold fronts which clear the air. Windy, turbulent days can create excessive image motion and cause difficulties in keeping the camera oriented for vertical photography, in staying on planned flight lines, and in maintaining constant flying heights.

The decision to fly or not to fly is one that must be made daily. The flight crew should be capable of interpreting weather conditions and of making sound decisions as to when satisfactory photography can be obtained. If possible, the flight crew should be based near the project so that they can observe the weather firsthand and quickly take advantage of satisfactory conditions.

18-9 Season of the Year

The season of the year is a limiting factor in aerial photography because it affects ground cover conditions and the sun's altitude. If photography is being taken for topographic mapping, the photos should be taken when the deciduous trees are bare, so that the ground is not obscured by leaves. In many places this occurs twice a year for short periods in the late fall and in early spring. Oak trees tend to hold many of their leaves until spring, when the buds swell and cause the leaves to fall. In areas with heavy oak cover, therefore, the most satisfactory period for aerial photography is that very short period in the spring between budding and leafing out. Sometimes aerial photography is taken for special forestry interpretation purposes, in which case it may be desirable for the trees to be in full leaf. Normally, aerial photography is not taken when the ground is snow-covered. Heavy snow not only obscures the ground but also causes difficulties in interpretation and in stereoviewing. Occasionally, however, a light snow cover can be helpful by making the ground surface more readily identifiable in tree-covered areas.

Another factor to be considered in planning aerial photography is the sun's altitude. Low sun angles produce long shadows, which can be objectionable because they obscure detail. Generally a sun angle of about 30° is the minimum acceptable for aerial photography. During the winter months of November through February, the sun never reaches a 30° altitude in some northern parts of the United States due to the sun's southerly declination. Aerial photography should therefore be avoided in those areas during these months, if possible. Often, snow cover will prevent photography during these periods anyway. For the other months, photography should be exposed during the middle portion of the day after the sun rises above 30° and before it falls below that altitude. For certain purposes, shadows may be desirable, since they aid in identifying objects. Shadows of trees, for example, help to identify the species. Shadows may also be helpful in locating photo-identifiable features such as fenceposts and power poles to serve as photo control points.

18-10 Flight Map

A flight map, as shown in Fig. 18-9, gives the project boundaries and flight lines the pilot must fly to obtain the desired coverage. The flight map is prepared on some existing map which shows the project area. United States Geological Survey quadrangle maps are frequently used. The flight map may also be prepared on small-scale photographs of the area, if they are available. In executing the planned photographic mission, the pilot finds two or more features on each flight line which can be identified both on the flight map and on the ground. The aircraft is flown so that lines of flight pass over the ground points. Alternatively, an airborne GPS receiver can be employed to guide the aircraft along predefined flight lines.

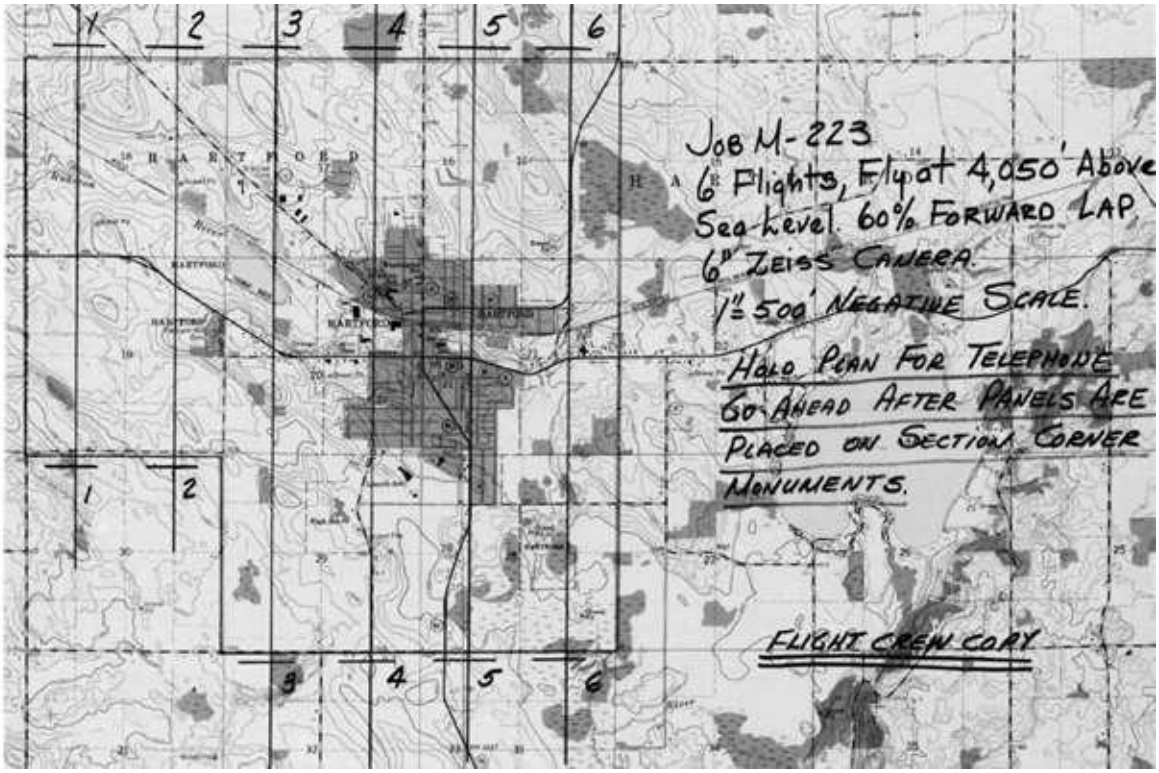


FIGURE 18-9 Example of a flight plan. (*Courtesy Ayres Associates, Inc.*)

Rectangular project areas are most conveniently covered with flight lines oriented north and south or east and west. As illustrated in Fig. 18-9, this is desirable because the pilot can take advantage of section lines and roads running in the cardinal directions and fly parallel to them.

If the project area is irregular in shape or if it is long and narrow and skewed to cardinal directions, it may not be economical to fly north and south or east and west. In planning coverage for such irregular areas, it may be most economical to align flight lines parallel to project boundaries as nearly as possible. Flight planning templates are useful for determining the best and most economical photographic coverage for mapping, especially for small areas. These templates, which show blocks of neat models, are prepared on transparent plastic sheets at scales that correspond to the scales of the base maps upon which the flight plan is prepared. The templates are then simply superimposed on the map over the project area and oriented in the position which yields best coverage with the fewest neat models. Such a template is shown in Fig. 18-10. The crosses represent exposure stations, and these may be individually marked on the flight map. This template method of flight planning is exceptionally useful in planning exposure station locations when artificial targets are used (see Sec. 16-2).

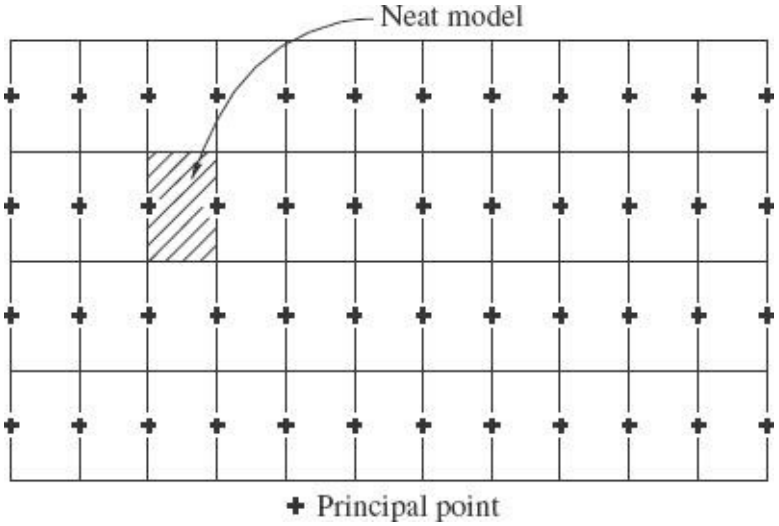


FIGURE 18-10 Transparent template of neat models used for planning aerial photography.

Once the camera focal length, photo scale, end lap, and side lap have been selected, the flight map can be prepared. The following example illustrates flight map preparation for a rectangular project area.

Example 18-8

A project area is 10 mi (16 km) long in the east-west direction and 6.5 mi (10.5 km) wide in the north-south direction (see Fig. 18-11). It is to be covered with vertical aerial photography having a scale of 1:12,000. End lap and side lap are to be 60 and 30 percent, respectively. A 6-in- (152.4-mm-) focal-length camera with a 9-in- (23-cm-) square format is to be used. Prepare the flight map on a base map whose scale is 1:24,000, and compute the total number of photographs necessary for the project.

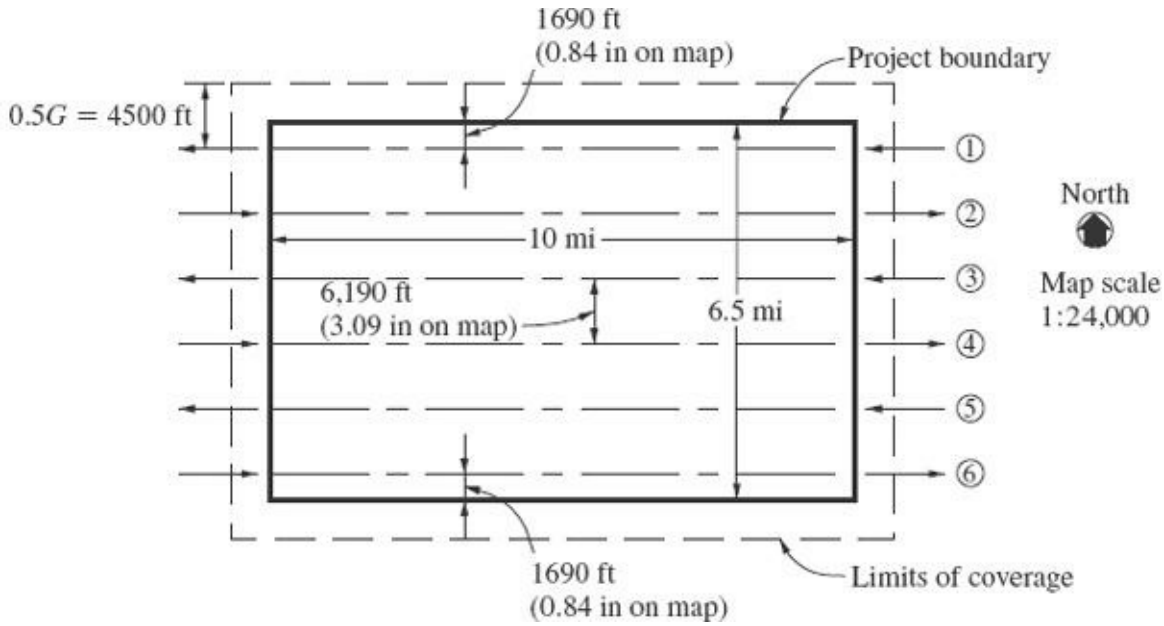


FIGURE 18-11 Project area for Example 18-9.

Solution

1. Fly east-west to reduce the number of flight lines.

2. Dimension of square ground coverage per photograph [photo scale = 1:12,000 (1 in/1000 ft)] is

$$G = 9 \text{ in} \times 1000 \text{ ft/in} = 9000 \text{ ft } (2800 \text{ m})$$

3. Lateral advance per strip (at 30 percent side lap) is

$$W = 0.7G = (0.7)(9000 \text{ ft}) = 6300 \text{ ft } (1900 \text{ m})$$

4. Number of flight lines. (Align the first and last lines with 0.3*G* (side-lap dimension) coverage outside the north and south project boundary lines, as shown in Fig. 18-11. This ensures lateral coverage outside of the project area.)

   Distance of first and last flight lines inside their respective north and south project boundaries (see Fig. 18-11) is

$$0.5G - 0.3G = 0.2G = 0.2(9000) = 1800 \text{ ft } (550 \text{ m})$$

   Number of spaces between flight lines:

$$\frac{6.5 \text{ mi} \times 5280 \text{ ft/mi} - 2 \times 1800 \text{ ft}}{6300 \text{ ft}} = 4.9 \text{ (round up to 5)}$$

   Number of flight lines = number of spaces + 1 = 6

5. Adjust the percent side lap and flight line spacing.

   Adjusted percent side lap for integral number of flight lines (include portion extended outside north and south boundaries):

$$2\left(0.5 - \frac{\text{PS}}{100}\right)G + (\text{no. spaces})\left(1 - \frac{\text{PS}}{100}\right)G = \text{total width}$$

$$2\left(0.5 - \frac{\text{PS}}{100}\right)9000 \text{ ft} + 5\left(1 - \frac{\text{PS}}{100}\right)9000 \text{ ft} = 6.5 \text{ mi} \times 5280 \text{ ft/mi}$$

$$2\left(0.5 - \frac{\text{PS}}{100}\right) + 5\left(1 - \frac{\text{PS}}{100}\right) = 3.813$$

$$\text{PS} = 31.2\% \quad \text{(Note slight increase.)}$$

   Adjusted spacing W_a between flight lines for integral number of flight lines:

$$W_a = \left(1 - \frac{31.2}{100}\right)G = 6190 \text{ ft } (1890 \text{ m})$$

6. Linear advance per photo (air base at 60 percent end lap):

$$B = 0.4G = (0.4)(9000 \text{ ft}) = 3600 \text{ ft } (1100 \text{ m})$$

7. Number of photos per strip (take two extra photos beyond the project boundary at both ends of each strip to ensure complete stereoscopic coverage):

$$\text{No. photos per strip} = \frac{10\text{ mi} \times 5280\text{ ft/mi}}{3600\text{ ft}} + 1 + 2 + 2$$

$$= 19.7\,(\text{use } 20)$$

8. Total number of photos:

$$20\text{ photos/strip} \times 6\text{ strips} = 120$$

9. Spacing of flight lines on the map:

$$\text{Map scale} = 1{:}24{,}000\,(1\text{ in} = 2000\text{ ft})$$

$$W_M = \frac{6190\text{ ft per strip}}{2000\text{ ft/in}} = 3.09\text{ in } (78.6\text{ mm})$$

Draw the flight lines at 3.09-in spacing on the map, with the first and last lines

[(0.5 – 31.2/100)9000 ft]/2000 ft/in = 0.84 in inside the project boundaries. ▲

Computer programs are now available for preparing flight plans. Figure 18-12 illustrates a flight plan for a highway corridor being prepared with the aid of a computer. Design variables including camera focal length, photo scale, end lap, and side lap are input to the computer. Base maps upon which the flight maps will be prepared can either be scanned into the computer from existing hard copies, or they can be downloaded from existing databases of topographic maps. Coordinates of points that delineate the boundaries of the area to be photographed must be input, and the coordinates of at least two control points must also be entered and their locations identified within the map area. The computer can then make all the same calculations that were demonstrated in Example 18-8 and prepare flight maps with superimposed flight lines. In addition to software devoted soley to preparing flight plans, there are various kinds of mapping software that allows users to easily develop their own flight plans by superimposing lines and points over existing imagery. In the most modern aerial photography and navigation systems, after designing the flight map, the computer determines the coordinates of the ends of the flight lines. Then the aircraft's navigation system, aided by an onboard GPS receiver (see Secs. 16-6 through 16-8), automatically guides the aircraft along the desired flight lines at the required altitude and exposes the photographs according to the given percent end lap.

FIGURE 18-12 Flight plan prepared on a computer. (*Courtesy University of Florida*)

18-11 Specifications

Most flight plans include a set of detailed specifications which outline the materials, equipment, and procedures to be used on the project. These specifications include requirements and tolerances pertaining to photographic scale (including camera focal length and flying height), end lap, side lap, tilt, crab, and photographic quality. The following is a sample set of detailed specifications for aerial photography (courtesy Ayres Associates, Inc.).

1. *General*. The engineer shall perform the necessary flying and photography to provide photographic coverage of an area approximately 8 square miles in extent shown on the sketch map attached hereto as exhibit A. The engineer may sublet this phase of the work to a qualified and experienced aerial photographic firm. The city, however, retains the right to approve or reject any or all such firms which the engineer may wish to engage.
2. *Scale*. Flight height above average ground shall be such that the negatives will have an average scale of 1 in = 500 ft (1:6000). Negatives having a departure from the specified scale by more than 5 percent because of tilt or abrupt changes in flying altitude must be corrected. The photographs shall be suitable for the compilation of the topographic maps specified herein, and the mapping flight height shall not vary from 3000 ft above mean terrain by more than 5 percent.
3. *End lap and side lap*. End lap shall be sufficient to provide full stereoscopic coverage of the area to be mapped. End lap shall average 63 percent, plus or minus 5 percent. End lap of less than 58 percent or more than 68 percent in one or more negatives shall be cause for rejection of the negatives in which such deficiency or excess occurs; unless within a stereoscopic pair, end lap exceeding 68 percent is necessary in areas of low elevation to attain the minimum 58 percent end lap in adjacent areas of high elevation. Wherever there is a change in direction of the flight lines, vertical photography on the beginning of a forward section shall end-lap the photography of a back section by 100 percent. Any negatives having side lap of less than 20 percent or more than 55 percent may be rejected.
4. *Tilt*. Negatives made with the optical axis of the aerial camera in a vertical position are desired.

Tilt of any negative by more than 5°, an average tilt of more than 1° for the entire project, or tilt between any two successive negatives exceeding 4° may be cause of rejection.

5. *Crab*. Crab in excess of 3° may be cause of rejection of the flight line of negatives or portions thereof in which such crab occurs.
6. *Quality*. The photographs shall be clear and sharp in detail and of uniform average density. They shall be free from clouds, cloud shadows, light streaks, static marks, or other blemishes which would interfere with their intended use. All photography shall be taken when the area to be mapped is free of snow, before foliation, and at such time as to ensure a minimum solar angle of 30°, except upon written authorization to the contrary by the city.
7. *Camera*. For topographic and contour mapping, photographs shall be exposed with a distortion-free 6-in- (152-mm-) focal-length precision aerial mapping camera equipped with a between-the-lens element shutter to produce negatives 9 in × 9 in (23 cm × 23 cm). The engineer shall furnish the city with a precision camera calibration report for the camera to be used.
8. *Contact prints*. The contact prints from the vertical negatives shall be printed on double-weight semimatte paper of suitable contrast.
9. *Photo index*. Photo indices shall be prepared by directly photographing, on safety base film at a convenient scale, the assembly of contact prints from all indexed and evaluated prints used. The photo index shall carry a suitable title, scale, and north point.
10. *Ownership of negatives*. All negatives shall become the property of the city and shall be delivered to the city upon completion of this contract, or may be stored indefinitely in the film library of the engineer at no added charge.

18-12 Cost Estimating and Scheduling

Cost estimating is an area of critical concern in the operation of any photogrammetric business, because if projects are let to contract that are underestimated, devastating financial results can ensue. In general, the items that must be considered in a cost analysis include material, labor, and overhead. In addition, a reasonable allowance for profit must be included.

Material costs are directly related to the quantity of each photogrammetric product to be prepared, and the procedures for calculating these quantities are quite straightforward. Example 18-8 illustrated the method of estimating the total number of aerial photos needed to cover a project area. Similar procedures are used to estimate other materials, and thus these costs can usually be estimated with fair accuracy. Overhead costs, which consist of salaries of administrative personnel, office and laboratory rental, electricity, water, heat, telephone, miscellaneous office supplies, etc., are also rather straightforward to determine. Labor costs, which generally constitute the major expense on photogrammetric projects, are considerably more difficult to estimate accurately, and this presents the greatest challenge to the estimator.

Although estimates can be made of the number of hours needed to perform each task involved in photogrammetric procedures, there are many unforeseen circumstances that can cause the actual time expended to deviate significantly from the estimates. The most realistic approach to estimating labor, therefore, is to rely on past experiences with projects of a similar nature. Obviously, it then becomes very important to keep detailed records of the actual costs incurred on the individual items of all projects. Because of significant variations of project complexities and rapidly changing cost factors, past records alone cannot be relied upon completely, and a good deal of intuition and subjective judgment is also necessary. This can normally be obtained only through years of experience.

In estimating, it is easy to omit small items, but enough of these over time can accumulate to cause a significant loss of revenue. Therefore, care must be exercised to prevent these omissions, and

the use of checklists is a good way to handle this problem.

Once the total number of labor hours has been estimated for each phase of a project, schedules for completion of the various operations can be planned on the basis of the number of instruments and personnel available to do the work. In addition to these factors, however, another important consideration is the amount of other work in progress and its status in relation to required completion dates. To arrive at realistic schedules, additional time in excess of that actually needed to perform the work must be added to account for uncontrollable circumstances. As an example, schedules for aerial photography and ground control surveys must account for possible delays due to inclement weather.

Every reasonable attempt should be made to accommodate clients with stringent scheduling needs. In some cases, to meet critical new scheduling requirements and still adhere to delivery dates already agreed upon, it may be necessary to consider hiring additional staff and running more than one work shift. Of course, the possibility of purchasing additional equipment also exists, but this should be done with caution and only when anticipated quantities of continued future work can justify the expenditures.

References

American Society of Photogrammetry: *Manual of Photogrammetry*, 5th ed., Bethesda, Md., 2004. chap. 9.

Graham, L. C.: "Flight Planning for Stereo Radar Mapping," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 9, 1975, p. 1131.

Hobbie, D.: "Orthophoto Project Planning," *Photogrammetric Engineering*, vol. 40, no. 8, 1974, p. 967.

Lafferty, M. E.: "Accuracy/Costs with Analytics," *Photogrammetric Engineering*, vol. 39, no. 5, 1973, p. 507.

Moffitt, F. H.: "Photogrammetric Mapping Standards," *Photogrammetric Engineering and Remote Sensing*, vol. 45, no. 12, 1979, p. 1637.

Paterson, G. L.: "Photogrammetric Costing," *Photogrammetric Engineering*, vol. 37, no. 12, 1971, p. 1267.

Ulliman, J. J.: "Cost of Aerial Photography," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 4, 1975, p. 491. U.S. Army Corps of Engineers: *Photogrammetric Mapping*, EM1110-1-1000, Available at: http://spatialdata.sam.usace.army.mil/organizations/survey/2002.

Walker, P. M., and D. T. Trexler: "Low Sun-Angle Photography," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 4, 1977, p. 493.

Wood, G.: "Photo and Flight Requirements for Orthophotography," *Photogrammetric Engineering*, vol. 38, no. 12, 1972, p. 1190.

Problems

18-1. The air base of a stereopair of vertical photos is 1400 m, and the flying height above average ground is 3000 m. If the camera has a 152-mm focal length and a 23-cm square format, what is the percent end lap?

18-2. Repeat Prob. 18-1, except that the focal length is 15,000 pixels and the format is 10,000 pixels squared.

18-3. Repeat Prob. 18-1, except that the air base is 4300 ft and flying height above average ground is 6500 ft.

18-4. For Prob. 18-1, if adjacent flight lines are spaced at 2260 m, what is the percent side lap?

18-5. For Prob. 18-2, if adjacent flight lines are spaced at 1390 m, what is the percent side lap?

18-6. For Prob. 18-3, if adjacent flight lines are spaced at 6300 ft, what is the percent side lap?

18-7. An average photo scale of 1:20,000 is required of vertical photos. What air base is required to achieve 60 percent end lap if the camera has a 6 in focal length and a 23-cm square format?

18-8. Repeat Prob. 18-7, except that required photo scale is 1:8000 and average end lap must be 55 percent.

18-9. Vertical photographs are to be exposed from 2300 m above average ground. If a *B/H′* ratio of 0.60 is required, what should be the length of the air base? What will the percent end lap be for these photos if the camera focal length is 152 mm and the format is 23 cm square?

18-10. Repeat Prob. 18-9, except that the photos were exposed from 4900 ft above ground and the required *B/H′* ratio is 0.65.

18-11. What is the *B/H′* ratio for vertical photography exposed with 55 percent end lap using a camera having a 152-mm focal length and a 9 in square format?

18-12. Repeat Prob. 18-11, except that end lap is 60 percent and camera focal length is 210 mm.

18-13. A project requires counting the number of people on a beach by using aerial photography. Assuming a 2-ft-diameter circle as a reasonable size for a person when viewed from above, and assuming a film with 80 line pairs per millimeter resolution will be used, what photo scale will be required? If a 152-mm-focal-length camera will be used, what is the required flying height above the beach?

18-14. A map with a scale of 1:8000 is to be compiled from vertical aerial photographs. The enlargement ratio from photo scale to map scale will be 5, and a 152-mm-focal-length camera will be used. What should be the flying height above average ground for the photography?

18-15. Repeat Prob. 18-14, except that a map with a scale of 1:6000 will be compiled and an enlargement ratio of 7 will be applied.

18-16. A *C* factor of 1500 will be applied in compiling a map having a contour interval of 3 m. What maximum flying height is acceptable, and what is corresponding photo scale if the camera has a 152-mm focal length?

18-17. Repeat Prob. 18-16 except that the contour interval is 5 ft and a *C* factor of 1300 will be applied.

18-18. An engineering design map is to be compiled from aerial photography. The map is to have a scale of 1:2400 and a 2-m contour interval. The enlargement ratio from photo scale to map scale is 5,

and the *C* factor is 1500. If the camera focal length is 152 mm, what is the required flying height above average ground, based upon required map scale? Based upon contour interval? Which condition controls flying height?

18-19. Repeat Prob. 18-18, except that map scale is 500 ft/in, the contour interval is 10 ft, and the *C* factor and enlargement ratio to be applied are 1500 and 7, respectively.

18-20. Vertical aerial photographs are taken from a flying height of 3500 m above average ground using a camera with a 210-mm-focal-length lens and a 23-cm square format. End lap is 60 percent at average terrain elevation. How many acres of ground are covered in a single photograph? In the neat model? (Assume 30 percent side lap.)

18-21. For Prob. 18-20, if low, average, and high terrain is 500, 600, and 700 m, respectively, above datum, what is the percent end lap at low terrain? At high terrain? What is the percent side lap at low terrain? At high terrain?

18-22. A rectangular area 10 mi in the north-south direction by 40 mi in the east-west direction is to be covered with aerial photography having a scale of 1:5000. End lap and side lap are to be 60 and 30 percent, respectively. A camera having a 23-cm square format is to be used. Compute the total number of photographs in the project, assuming that the flight strips are flown in an east-west direction and that the coverage of the first and last flight lines is 75 percent within the project boundary. Also add two photos at the ends of each strip to ensure complete coverage.

18-23. If a flight map is to be prepared for Prob. 18-22 on a base map having a scale of 1:24,000, what should be the spacing of flight lines on the map? What is the map distance between successive exposures along a flight line?

18-24. A transparent template of neat models, similar to that shown in Fig. 18-10, is to be prepared to overlay on a map having a scale of 1:12,000. What should be the dimensions of neat models on the template if the camera format is 23 cm square, photo scale is 1:6000, end lap is 60 percent, and side lap is 30 percent?

18-25. Repeat Prob. 18-24, except that map scale is 1:120,000 and photo scale is 1:24,000.

18-26. A rectangular project area 3 km in the north-south direction and 4 km in the east-west direction is to be photographed at a scale of 1:4000. End lap and side lap are to be 60 and 30 percent, respectively, and the camera format is 23 cm square. Compute the total number of photographs needed to cover this area, assuming that flight lines will run east-west and that the first and last lines will be flown so that the adjusted side lap will extend outside the project boundaries. Add two photos at the ends of each strip to ensure complete coverage. Prepare a flight map showing the flight lines, assuming the base map has a scale of 1:24,000.

---

[1] Although a significant number of projects involve terrestrial or close-range photogrammetry, in this chapter aerial photogrammetry is assumed.

[2] The complete set of National Map Accuracy Standards is available at the following website: http://rmmcweb.cr.usgs.gov/public/nmpstds/nmas647.html

CHAPTER 19
Terrestrial and Close-Range Photogrammetry

19-1 Introduction

Terrestrial photogrammetry is an important branch of the science of photogrammetry. It deals with photographs taken with cameras located on the surface of the earth. The cameras may be handheld, mounted on tripods, or suspended from towers or other specially designed mounts. The term *close-range photogrammetry* is generally used for terrestrial photographs having object distances up to about 300 m. With terrestrial photography the cameras are usually accessible, so that direct measurements can be made to obtain exposure station positions, similar to airborne GPS control with aerial photography. With some terrestrial cameras, angular orientation can also be measured or set to fixed values, so that all elements of exterior orientation of a terrestrial photo are commonly known and need not be calculated. These known exterior orientation parameters are a source of control for terrestrial photos, replacing in whole or in part the necessity for locating control points in the object space.

Terrestrial photography may be *static* (photos of stationary objects) or *dynamic* (photos of moving objects). For static photography, slow, fine-grained, high-resolution films may be used and the pictures taken with long exposure times. Stereopairs can be obtained by using a single camera and making exposures at both ends of a baseline. In taking dynamic terrestrial photos, fast films and rapid shutter speeds are necessary. If stereopairs of dynamic occurrences are required, two cameras located at the ends of a baseline must make simultaneous exposures.

19-2 Applications of Terrestrial and Close-Range Photogrammetry

Historically the science of photogrammetry had its beginning with terrestrial photography, and topographic mapping was among its early applications. Terrestrial photos were found especially useful for mapping rugged terrain which was difficult to map by conventional field-surveying methods. Although it was known that topographic mapping could be done more conveniently using aerial photos, no practical method was available for taking aerial photographs until the airplane was invented. Following the invention of the airplane, emphasis in topographic mapping shifted from terrestrial to aerial methods. Terrestrial photogrammetry is still used in topographic mapping, but its application is usually limited to small areas and special situations such as deep gorges or rugged mountains that are difficult to map from aerial photography. Other topographic applications of terrestrial photogrammetry are in mapping construction sites, areas of excavation, borrow pits, material stockpiles, etc.

Through the years terrestrial photogrammetry has continued to gain prominence in numerous diversified nontopographic applications. Examples of nontopographic applications occur in such areas as aircraft manufacture, shipbuilding, telecommunications, robotics, forestry, archaeology, anthropology, architecture, geology, engineering, mining, nuclear industry, criminology, oceanography, medicine, dentistry, and many more. In the field of medicine, X-ray photogrammetry has been utilized advantageously for measuring sizes and shapes of body parts, recording tumor growth, studying the development of fetuses, locating foreign objects within the body, etc.

Figure 19-1 illustrates an application of close-range photogrammetry to antenna design. In this

figure, close-range photographs are taken of a parabolic radio antenna. Coordinates of points on the antenna were computed to verify its dimensional integrity of the rocket system. Figure 19-2 demonstrates the use of close-range photogrammetry to determine the dimensions of a ship's propeller. In this application, the propeller's size and shape were determined so that a replacement could be fabricated with the correct dimensions. Figure 19-3 illustrates the use of close-range photogrammetry to determine the shape of an inflatable antenna. This information was used to verify the size and shape of the antenna after inflation. Note that a projector is being used to produce the white dots needed to define the surface. In addition, the use of two synchronized cameras allows three-dimensional measurement of dynamic objects.

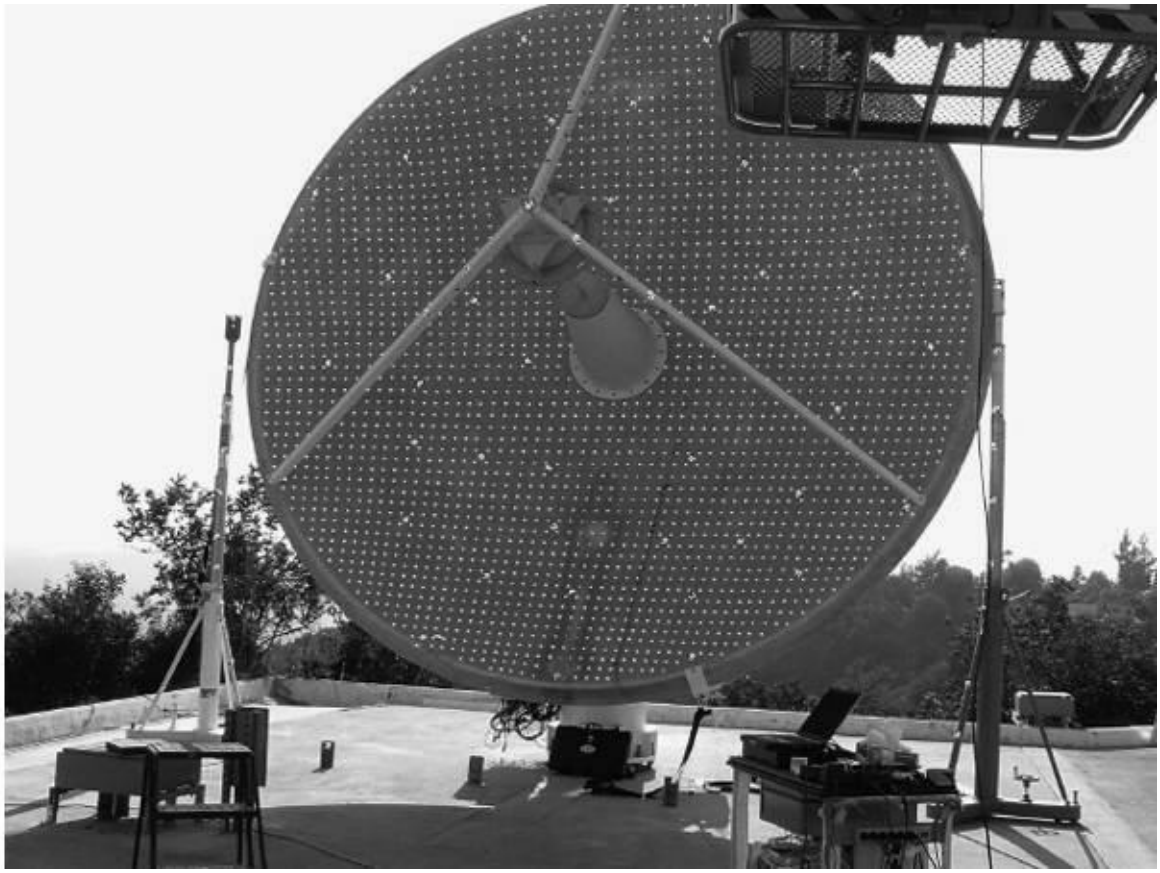


FIGURE 19-1 Application of close-range photogrammetry to the determination of the precise shape of a parabolic antenna. Note the artificial targets (white dots) placed on the structure. (*Courtesy Geodetic Systems, Inc.*)

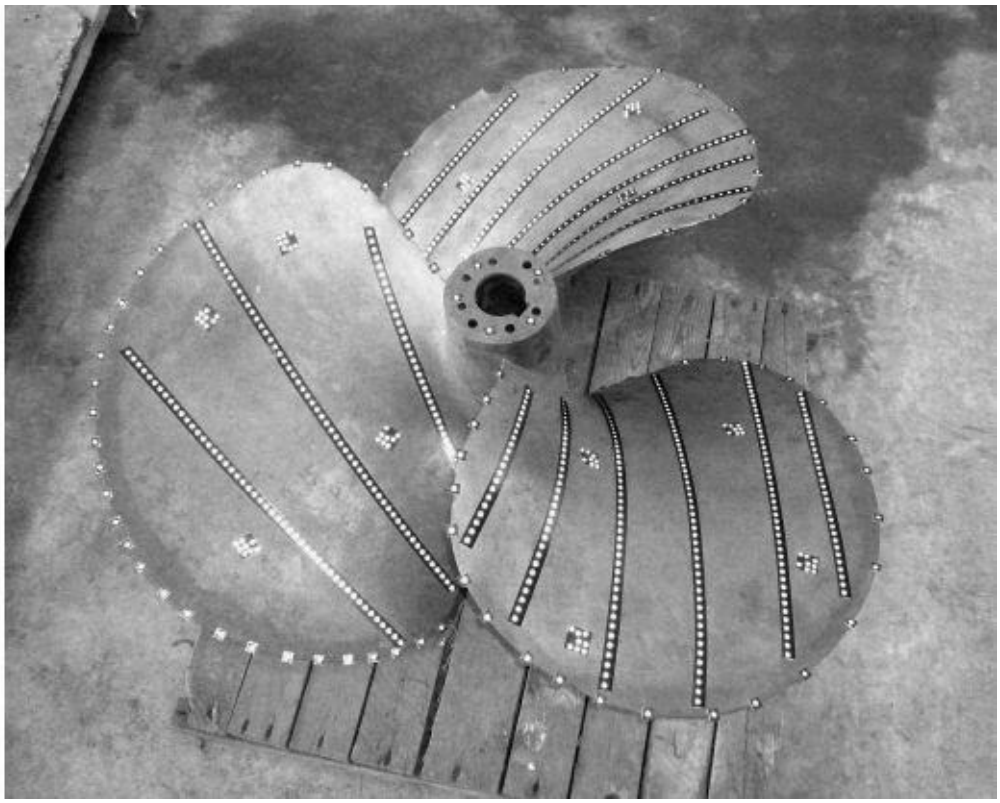


FIGURE 19-2 Application of close-range photogrammetry for determining the size and shape of a ship's propeller. Note the artificial targets (white dots) placed on the structure. (*Courtesy Geodetic Systems, Inc.*)

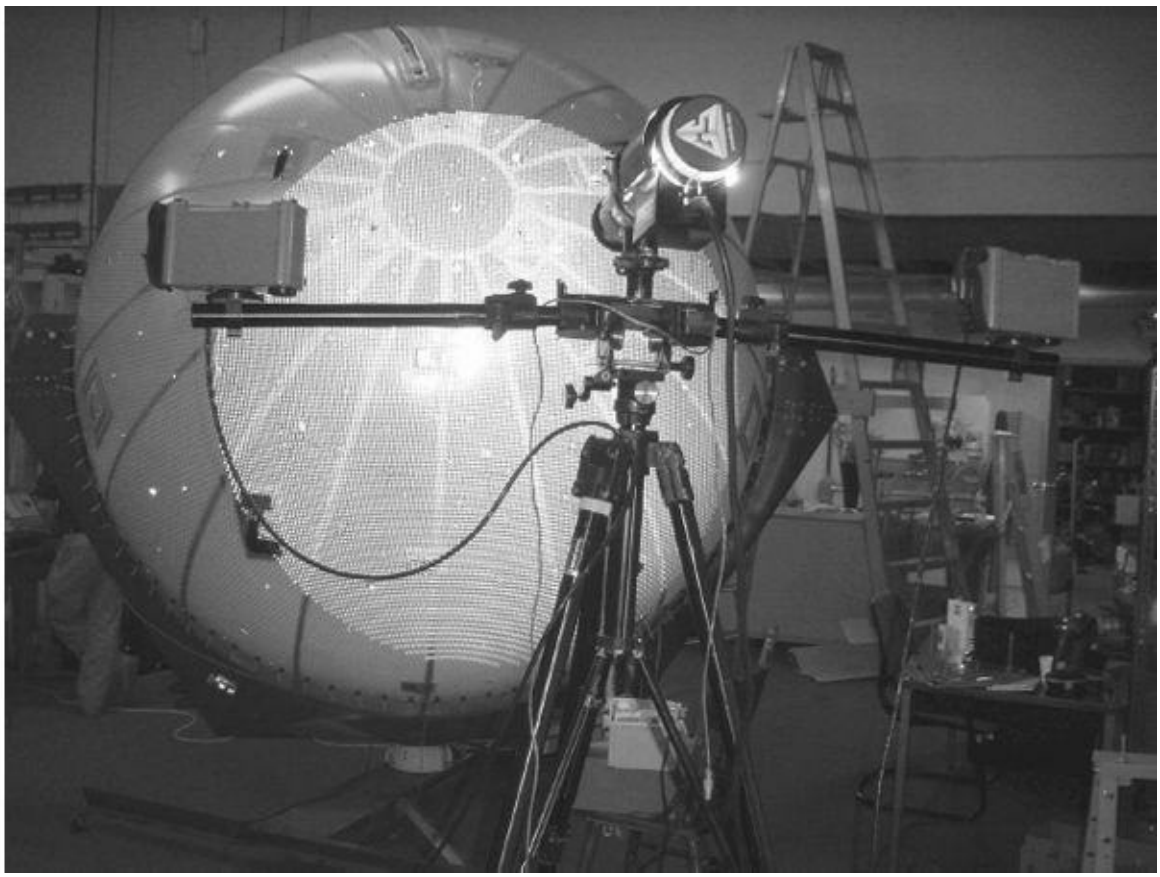


FIGURE 19-3 Application of close-range photogrammetry to determine the size and shape of an inflatable antenna. (*Courtesy Geodetic Systems, Inc.*)

Terrestrial photogrammetry has been used to great advantage as a reliable means of investigating traffic accidents. Photos that provide all information necessary to reconstruct the accident may be rapidly obtained. Time-consuming sketches and ground measurements, which all too often are erroneous, are not needed, and normal traffic flow can be restored more quickly. Terrestrial photogrammetry has been widely practiced in accident investigation for many years in several European countries.

Terrestrial photogrammetry has become a very useful tool in many areas of scientific and engineering research for several reasons. One reason is that it makes possible measurements of objects which are inaccessible for direct measurement. Also, measurements can be obtained without actually touching delicate objects. In some experiments, such as measurements of water waves and currents, physical contact during measurement would upset the experiment and render it inaccurate. Cameras which freeze the action at a particular instant of time make possible measurements of dynamic events such as deflections of beams under impact loads. Because of the many advantages and conveniences offered by terrestrial photogrammetry, its importance in the future seems assured.

19-3 Terrestrial Cameras

A variety of cameras are used in terrestrial photography. All fall into one of two general classifications: *metric* or *nonmetric*. The term *metric camera*, as used here, includes those cameras manufactured expressly for photogrammetric applications. They have fiducial marks built into their focal planes, or have calibrated CCD arrays, which enable accurate recovery of their principal points. Metric cameras are stably constructed and completely calibrated before use. Their calibration values for focal length, principal-point coordinates, and lens distortions can be applied with confidence over long periods.

The Dynamo camera of Fig. 19-4 is a digital camera that is generally used as one of a synchronized pair, which can capture simultaneous images necessary for dynamic applications.

FIGURE 19-4 Dynamo close-range camera (bottom) with strobe attachment (top). (*Courtesy Geodetic Systems, Inc.*)

The INCA3 (*In*telligent Digital *Ca*mera) of Fig. 19-5 is a digital terrestrial camera which is also used primarily for industrial applications. This camera uses a CCD array having a either a 2029 × 2044 or 3500 × 2300 pixel format, with each pixel having a 12-bit gray level. The camera's angular field of view is 77° × 56° with a 21 mm lens. Images are recorded on a flash memory card or directly transferred to a computer in real-time. Accuracies of better than 1 part in 200,000 of the object size can be achieved with this camera.

FIGURE 19-5 The INCA3 digital close-range camera. (*Courtesy Geodetic Systems, Inc.*)

Phototheodolites and *stereometric cameras* are two special types of terrestrial camera systems in the metric classification. A phototheodolite (see Fig. 1-1) is an instrument that incorporates a metric camera with a surveyor's theodolite. With this instrument, precise establishment of the direction of the optical axis can be made. A stereometric camera system consists of two identical metric cameras which are mounted at the ends of a bar of known length. The optical axes of the cameras are oriented perpendicular to the bar and parallel with each other. The length of the bar provides a known baseline length between the cameras, which is important for controlling scale.

Nonmetric cameras are manufactured for amateur or professional photography where pictorial quality is important but geometric accuracy requirements are generally not considered paramount. These cameras do not contain fiducial marks, but they can be modified to include them. Nonmetric cameras can be calibrated and used with satisfactory results for many terrestrial photogrammetric applications.

19-4 Matrix Equations for Analytical Self-Calibration

The camera calibration parameters, or *elements of interior orientation*, were described in Sec. 3-10. The effects of these distortion parameters must be accounted for in order to make precise photogrammetric measurements. When using metric cameras, this is not of great concern due to the stability of the parameters over long intervals of time—although they should be calibrated periodically based on the manufacturer's recommendation. On the other hand, nonmetric cameras cannot be expected to maintain stable calibration parameters even between sessions. In many cases, however, nonmetric cameras can provide very accurate results when using *analytical self-calibration*.

Analytical self-calibration is a computational process wherein camera calibration parameters are included as unknowns in the photogrammetric solution, generally in a combined interior-relative-absolute orientation referred to as a *self-calibrating bundle adjustment*. This process can be used for both aerial and terrestrial photos. In some circumstances, the data acquired during a mission can be used to calibrate the camera. This case ensures that the solved calibration parameters represent the interior orientation of the camera at the time of exposure of the photographs. Alternatively, a specially-designed calibration pattern with optimally configured three-dimensional control points can be used in conjunction with photography taken with optimally-oriented camera exposure stations (see Sec. 19-9). In this case, more precise estimates of the parameters can be obtained due to the favorable geometry. However, there is no guarantee that they will be the same when the camera is used during the mission due to internal movement of the mechanical elements of the camera in the time between calibration and acquisition.

The process of analytical self-calibration uses collinearity equations that have been augmented with additional terms to account for adjustment of the calibrated focal length, principal-point offsets, and symmetric radial and decentering lens distortions. In addition, the equations might include corrections for atmospheric refraction. The classic form of the augmented collinearity equations is shown in Eq. (19-1). Note that the model presented here is valid for both film and digital photos.

$$x_a = x_0 - \bar{x}_a\left(k_1 r_a^2 + k_2 r_a^4 + k_3 r_a^6\right) - \left(1 + p_3 r_a^2\right)\left[p_1\left(2\bar{x}_a^2 + r_a^2\right) + 2p_2 \bar{x}_a \bar{y}_a\right] - f\frac{r}{q}$$

$$y_a = y_0 - \bar{y}_a\left(k_1 r_a^2 + k_2 r_a^4 + k_3 r_a^6\right) - \left(1 + p_3 r_a^2\right)\left[2p_1 \bar{x}_a \bar{y}_a + p_2\left(2\bar{y}_a^2 + r_a^2\right)\right] - f\frac{s}{q} \qquad (19\text{-}1)$$

where

x_a, y_a = measured photo coordinates related to fiducials

x_0, y_0 = coordinates of the principal point

$\bar{x}_a = x_a - x_0$

$\bar{y}_a = y_a - y_0$

$r_a^2 = \bar{x}_a^2 + \bar{y}_a^2$

k_1, k_2, k_3 = symmetric radial lens distortion coefficients

p_1, p_2, p_3 = decentering distortion coefficients

f = calibrated focal length

r, s, q = collinearity equation terms as defined in Eqs. (D-11) and (D-12)

The interior orientation parameters x_0, y_0, f, k_1, k_2, k_3, p_1, p_2, and p_3, which appear in Eq. (19-1) are

included as unknowns in the solution, together with w, j, k, X_L, Y_L, and Z_L for each photo and X_A, Y_A, and Z_A for each object point. These equations are nonlinear, therefore Taylor's series is used to linearize them, and an iterative solution is made. The matrix form of the linearized augmented collinearity equations is shown in Eq. (19-2).

$$\dot{B}_{ij}\dot{\Delta}_i + \ddot{B}_{ij}\ddot{\Delta}_i + \dddot{B}_{ij}\dddot{\Delta} = \varepsilon_{ij} + V_{ij} \tag{19-2}$$

where

$$\dddot{B}_{ij} = \begin{bmatrix} b_{1,7} & b_{1,8} & b_{1,9} & b_{1,10} & b_{1,11} & b_{1,12} & b_{1,13} & b_{1,14} & b_{1,15} \\ b_{2,7} & b_{2,8} & b_{2,9} & b_{2,10} & b_{2,11} & b_{2,12} & b_{2,13} & b_{2,14} & b_{2,15} \end{bmatrix}$$

$$\dddot{\Delta} = \begin{bmatrix} d_{x_0} \\ d_{y_0} \\ d_{k_1} \\ d_{k_2} \\ d_{k_3} \\ d_{p_1} \\ d_{p_2} \\ d_{p_3} \\ d_f \end{bmatrix}$$

Matrix $\dddot{B}_{ij}$ contains the partial derivatives of Eq. (19-1) with respect to the camera calibration parameters evaluated at the initial approximations. Matrix $\dddot{\Delta}$ contains the corrections for the initial approximations of the camera calibration parameters. The matrices $\dot{B}_{ij}$, $\ddot{B}_{ij}$, $\dot{\Delta}_i$, $\ddot{\Delta}_i$, ε_{ij}, and V_{ij} are the same as those in Eq. (17-4), with the exception that matrices ε_{ij} and V_{ij} are computed using the augmented collinearity equations. Weighted observations of the calibration parameters can be included in a similar way as ground control points and aerial control are in Sec. 17-6, by including the observation equations shown in matrix form in Eq. (19-3).

$$\dddot{\Delta} = \dddot{C} + \dddot{V} \tag{19-3}$$

Normally, direct observations of the calibration parameters are not made; however, weighting provides a method of constraining them, which in some cases is necessary to produce a convergent solution. The partitioned normal equations for a bundle adjustment with analytical self-calibration are shown in matrix form in Eq. (19-4).

$$\begin{bmatrix} \dot{N} & \bar{N} & \tilde{N} \\ \bar{N}^T & \ddot{N} & \hat{N} \\ \tilde{N}^T & \hat{N}^T & \dddot{N} \end{bmatrix} \begin{bmatrix} \dot{\Delta} \\ \ddot{\Delta} \\ \dddot{\Delta} \end{bmatrix} = \begin{bmatrix} \dot{K} \\ \ddot{K} \\ \dddot{K} \end{bmatrix} \tag{19-4}$$

Equation (19-4) is the same as the partitioned normal equations shown in Eq. (17-26), except with the following added terms:

$$\tilde{N} = \begin{bmatrix} \tilde{N}_1 \\ \tilde{N}_2 \\ \tilde{N}_3 \\ \vdots \\ \tilde{N}_m \end{bmatrix} \quad \text{where} \quad \tilde{N}_i = \sum_{j=1}^{n} \dot{B}_{ij}^T W_{ij} \dddot{B}_{ij}$$

$$\hat{N} = \begin{bmatrix} \hat{N}_1 \\ \hat{N}_2 \\ \hat{N}_3 \\ \vdots \\ \hat{N}_n \end{bmatrix} \quad \text{where} \quad \hat{N}_j = \sum_{i=1}^{m} \ddot{B}_{ij}^T W_{ij} \dddot{B}_{ij}$$

$$\dddot{N} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \dddot{B}_{ij}^T W_{ij} \dddot{B}_{ij} \right) + \dddot{W}$$

$$\dddot{K} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \dddot{B}_{ij}^T W_{ij} \varepsilon_{ij} \right) + \dddot{W}\dddot{C}$$

With the inclusion of the extra unknowns, it follows that additional independent equations will be needed to obtain a solution. In addition, the numerical stability of analytical self-calibration is of serious concern. Merely including the additional parameters does not guarantee that they will be precisely resolved. It is necessary to have special constraints and/or geometric configurations to ensure a stable solution. For example, with nominally vertical aerial photography if the object points are at roughly the same elevation, then x_o, y_o, and f are strongly correlated with X_L, Y_L, and Z_L, respectively. Similarly, nominally horizontal terrestrial photography of a wall taken with the focal plane of the camera parallel to the surface of the wall results in f and x_o being correlated with X_L and Y_L, and y_o being correlated with Z_L. Given these correlations, the solution may not produce satisfactory results. This problem can be overcome or at least alleviated if there is significant variation in the

depth of the object field with respect to the camera, by using highly convergent photography, by making observations of the camera position and/or attitude, or by using photos with varying rotations about the optic axis of the camera. In addition, to recover the lens distortion parameters accurately, it is necessary to have many redundant object points whose images are well distributed across the format of the composite of all photographs. In other words, the combined images of objects points from all photographs should be distributed over the extents of the format. For example, one photo may have image points only on the left half of the format and another photo may have points only on the right half of the format, but taken as a composite, images have been distributed across the extents of the format.

19-5 Initial Approximations for Least Squares Adjustment

As mentioned in the previous section, initial approximations of all unknown parameters are required for a least squares adjustment when using linearized observation equations. Thus, one must have sufficient preliminary estimates of the position and angular orientation of the camera stations ω, φ, κ, X_L, Y_L, and Z_L), object space coordinates of all imaged points (X_A, Y_A, and Z_A), and values of all camera calibration parameters (x_0, y_0, f, k_1, k_2, k_3, p_1, p_2, and p_3) prior to implementation. This section will focus on the first two sets, with the initialization of camera calibration parameters addressed in Sec. 19-6. Obtaining initial approximations for terrestrial photography can be much more difficult than for aerial photography. This not only stems from the inherent differences in geometry of the configurations, but also from the fact that in many close range and terrestrial applications, the photography was taken without the intention of using it for photogrammetry (e.g., historical photos and photos used for accident reconstruction). In these instances, the photographer likely did not note the position and angular orientation of the camera during exposure. There are, however, both manual and automatic methods for obtaining initial approximations for terrestrial and close range photogrammetric adjustments.

Manual methods can often provide adequate initial approximations when automation is not an option. Perhaps the most complex unknowns to estimate manually prior to adjustment using terrestrial/close range photography are the exterior orientation angles ω, φ, and κ. For near-vertical aerial photos, these angles can easily be estimated with sufficient accuracy by using $\omega = \varphi = 0°$, and estimating κ using the direction of flight or via the method described in Sec. 11-6. In terrestrial photography, however, there is no simple method for their estimation due to the non-intuitive values of ω, φ, and κ. For example, a photo taken while tilted 70° from vertical with 180° swing (i.e., the y-axis is up in the photo), and pointed with an azimuth of 83° is oriented with $\omega = 26.05°$, $\varphi = –74.379°$, and $\kappa = –63.09°$. Under the same configuration, except pointed with an azimuth of 173°, the exterior orientation angles are $\omega = –75.90°$, $\varphi = –6.79°$, and $\kappa = –178.30°$. This shows that simply changing the pointing direction of the camera affects ω, φ, and κ in ways that are not easily visualized. This example also illustrates a method for estimating the exterior orientation angles manually. First, they are estimated using the tilt-swing-azimuth system described in Sec. 10-3 and then they are converted using the method in Sec. D-10 to obtain ω, φ, and κ. In fact, a bundle adjustment solution can be implemented using tilt, swing, and azimuth. However, photogrammetric software typically uses the omega-phi-kappa convention. As for the position of the camera, one can use the conditions of acquisition and scene in a photo to estimate the position from where it was taken. For example, if a person standing and holding the camera naturally took the photo, one can estimate the height of the camera above the ground as being slightly less than the height of the photographer. Estimates of object space coordinates for exposure stations and tie points can be deduced by approximating their

positions relative to points with known coordinates and their apparent positions based on the scenes. That is, control points can be used to determine the orientation and scale of the object space imaged in the photo. These can then be used to estimate the distance and direction from the control points to the tie points and exposure stations yielding their object space coordinates. This procedure requires practice to become proficient at rapidly and reliably obtaining these approximations.

Although manual methods are often reliable for small projects when performed with care, automatic methods are preferred when there are many photos involved. There are several methods for automatic initialization, only two of which are presented here. One method begins by using space resections (see Sec. 11-6) to solve for the exterior orientation parameters of two images. The exterior orientation parameters are then used to find the object space coordinates of tie points in the two images via space intersection (see Sec. 11-7). The object space coordinates of these tie points are used to perform space resection for other photos in the project. The resulting exterior orientation parameters can then be used in subsequent space intersections. The process is repeated until initial approximations for all images and tie points are found. Note that this method requires at least four control points in at least two images for a least squares solution of the first space resections. In industrial applications, which are usually performed in laboratory environments with metric cameras, automatic initialization is often achieved using *exterior orientation devices*. These devices consist of a single apparatus with several targets on it that are used to define the object space coordinate system. Figure 19-6 shows a five-target exterior orientation device. The relative positions of the targets on the device are known to a very high degree of precision, and serve as control points for the initial space resections.

FIGURE 19-6 The AutoBar exterior orientation device (*Courtesy Geodetic Systems, Inc.*)

Another method, similar to the previous one, utilizes relative orientation (see Sec. 11-10), and does not require that the minimum number of control points be visible in a single photo. This method begins by first performing relative orientation between the pair of photos with the most favorable geometry (see Sec. 19-9). Next, space intersection is used to find the model space coordinates of all tie points in the two images. These model space coordinates are used in a space resection to find the exterior orientation parameters—relative to the model—of the third photo. The model space coordinates of the remaining tie points in the third image not used for space resection are found through space intersection. The process is repeated on all subsequent images until all the model space exterior orientation parameters and coordinates of all tie points are found. Following this, a three dimensional conformal coordinate transformation from the model space to the (ground) control

system is solved using control points and applied to the model space coordinates of the tie points and exposure stations, yielding their estimates. Finally, the initial approximations for the exterior orientation angles can be found using the method described in Sec. 17-7.

If collinearity is used in the analytical relative orientation, initial approximations are needed for the relative orientation parameters and the coordinates of the tie points. However, the use of coplanarity (Sec. 11-5) to perform analytical relative orientation can circumvent the need for initial approximations of tie point model coordinates since they are not used in that method, although approximations of the five relative exterior orientation parameters are still required. Also, note that initial approximations are also required for space resection and space intersection. There are known, reliable methods for automatically finding initial approximations for space resection and space intersection solutions, which can be found in the references at the end of this chapter. However, relative orientation continues to be an active area of research with multiple *direct* and *search-based* methods described. These methods are beyond the scope of this book and the reader is encouraged to consult the reference section for information on them.

19-6 Solution Approach for Self-Calibration Adjustment

Carrying out a self-calibrating bundle adjustment of terrestrial/close range photography can be a complicated task. This section presents a method for adjustment that aids in troubleshooting and helps ensure convergence. The approach is designed for a bundle adjustment program that has input consisting of pass point measurements, control point measurements, initial approximations, and a priori standard deviations of all observations and camera calibration parameters. Typically, there are no direct observations of the camera calibration parameters. However, treating them as observations allows one to constrain them in the adjustment by weighting. Approximations for the focal length, f, can be determined using the camera manufacturer's specifications. The remaining calibration parameters can be initialized at zero. Before describing the approach, it should be stated that it is not uncommon for some calibration parameters to be too sensitive to calculate. On a related note, the guidelines given in Sec. 19-9 should be followed as much as possible before attempting an adjustment. Also, note that many of the concepts and procedures in Sec. 17-9 are applicable to terrestrial/self-calibration adjustments.

The first step of the approach is to adjust two photos. The photos selected for this initial adjustment should be convergent with a large amount of overlap between them and a sufficient amount of control points. Starting with two photos instead of attempting to adjust the entire set of photos reduces the chances of having multiple blunders, and therefore simplifies troubleshooting. One can attempt to adjust the two images with the calibration parameters "loosened," giving them high a priori standard deviations, and thus allowing them to be solved in the adjustment. However, this often leads to divergence due to a combination of insufficient distribution of tie points throughout the photos, homogeneous depth of field, inadequate redundancy, and poor geometric configuration (non-convergence) of the photos. In this case, the calibration parameters should be constrained by assigning them very small a priori standard deviations. If the adjustment fails to converge with constrained parameters, check for blunders and make sure initial approximations are consistent.

Once the two-photo adjustment converges, check the pass point residuals to identify any blunders, remeasuring them if found. Add photos one at a time, rerunning the adjustment and again checking for blunders with each addition. Previously estimated initial approximations can be updated using their solved values after each adjustment. This will speed up subsequent solutions by reducing the number of iterations required. After many photos have been added to the adjustment, the calibration parameters can be loosened.

If the adjustment diverges when the calibration parameters are loosened, it is advisable to loosen only f and k_1 since they are almost always significant. If this solution converges, one can then attempt to loosen x_0, y_0, and k_2, which are often significant. The remaining parameters require high redundancy and very strong geometry for them to be resolved. Due to their high correlation, it may be advisable to constrain p_1 and p_2 when x_0 and y_0 are loosened or vice-versa. One should also check the significance of each of the calibration parameters after the adjustment. This can be done by using a *t-test* of the significance of the adjusted parameter's departure from its initial approximation. The formula for the t-statistic is in Eq. (19-5), where $\hat{a}$ is one of the estimated calibration parameters (x_0, y_0, f, k_1, k_2, k_3, p_1, p_2, or p_3), a_0, is its initial approximation, and $s_{\hat{a}}$ is the estimated parameter's standard error which can be calculated using Eq. (17-16).

$$t_x = \frac{a_0 - \hat{a}}{s_{\hat{a}}} \tag{19-5}$$

If a calibration parameter is not significantly different from its initial approximation given its degrees of freedom, it should be excluded from the adjustment.

As mentioned in Sec. 17-6, the standard error of unit weight, S_0, should be close to one if all observations are properly weighted. If S_0 is too high, then estimated a priori standard deviations are too low and should be increased. If S_0 is too high, then estimated a priori standard deviations are too high, and should be decreased. A chi-squared can be used to test whether or not S_0 is significantly different than 1 by using the formula in Eq. (19-6). In Eq. (19-6), r is the degrees of freedom, and $\sigma^2 = 1$.

$$\chi^2 = \frac{rS_0^2}{\sigma^2} \tag{19-6}$$

Description of the t-test, chi-squared test, and their associated tables can be found in most introductory statistics textbooks.

19-7 Control for Terrestrial Photogrammetry

In terrestrial photogrammetry, the object space is often relatively close to the camera; in fact, in many cases the photographs are taken in laboratories. Object distances vary from a few centimeters up to 300 m or more, and the objects photographed vary in size from articles as small as human teeth and smaller, to very large buildings or ships. In any case, if accurate maps are to be made of photographed objects, control will be required.

In terrestrial photogrammetry there are basically four different methods of establishing control: (1) imposing the control on the camera by measuring its position and orientation with respect to a coordinate system or with respect to the photographed object, (2) locating control points in the object space in a manner similar to locating control for aerial photography, (3) combining camera control and object space control points, and (4) using a free-network adjustment with scale control only.

In the first method, no control points need appear in the object space. Rather, the position and orientation of the camera or cameras are measured with respect to the object itself. If a planar object is being photographed from a single camera station, control requirements may be satisfied by

measuring the distance from the camera to the plane surface and orienting the camera optical axis perpendicular to the surface. Perpendicular orientation can be accomplished by mounting a plane-surfaced mirror parallel to the object plane and then moving the camera about until the reflection of the camera lens occupies the center of the field of view. If the camera focal length is known, a complete planimetric survey of the object can then be made.

If stereopairs of photos are taken, the control survey can consist of measuring the horizontal distance and difference in elevation between the two camera stations and also determining the orientations of the camera optical axis for each photo. Phototheodolites enable a complete determination of camera orientation and direction of optical axis. Stereometric cameras automatically provide control by virtue of their known baseline length and parallel optical axes. In exposing stereopairs with less elaborate cameras, horizontal orientation can be enforced by using level vials or tilt sensors, and parallel orientation of the camera axes can be accomplished by reflection from parallel mirrors.

In the second method of controlling terrestrial photos, points should be selected in the object space which provide sharp and distinct images in favorable locations in the photographs. Their positions in the object space should then be carefully measured. If no satisfactory natural points can be found in the object space, artificial targets may be required. Targets should be designed so that their images appear sharp and distinct in the photos. White crosses on black cards may prove satisfactory. If the object space is small and the control points are close together, measurements for locating the targets may be made directly by means of graduated scales. If the object space is quite large or if the control points are inaccessible for direct measurement, triangulation with precise theodolites set up at the ends of a carefully measured baseline may be necessary. In some cases a premeasured grid pattern may be placed in the object space and photographed along with the object, thereby affording control.

If the object being photographed is stationary, control points may be located on the object. Corners of window frames, for example, may be used if a building is being photographed. If a dynamic event is being photographed at increments of time, for example, photographing beam deflections under various loads, then targets may have to be mounted on some stationary framework apart from the object. By means of surveyor's levels, targets may be set at equal elevations, thereby providing a horizontal line in the object space. Vertical lines may be easily established by hanging plumb bobs in the object space and attaching targets to the string.

The third method of controlling terrestrial photography is a combination of the first two methods. This third approach is generally regarded as prudent because it provides redundancy in the control, which prevents mistakes from going undetected and also enables increased accuracy to be obtained.

The fourth control method uses an arbitrary coordinate system, with the scale of the model being defined through one or more known lengths that appear in the object space. The known lengths may be based upon distance measurements made between target points on the object. In other cases, *scale bars* may be used. A scale bar is a long item such as a metal rod which has a known length. By placing one or more scale bars in proximity to the object, photo coordinate measurements can be made on the images of the ends of the bar, thus including the scale bar in the three-dimensional model of the object space. By constraining the distances between these endpoints to their known values, the scale of the overall object will be established. The arbitrary coordinate system can be established by either setting the position and angular attitude of one of the exposures to some nominal set of values (say, $\omega = \phi = k = X_L = Y_L = Z_L = 0$) or through the use of an exterior orientation device as described in Sec. 19-5. The remaining exposures and/or object points will then be determined relative to this set of defined values. After all coordinates have been obtained, a three-dimensional conformal coordinate transformation (see Sec. C-7) can be used to relate the arbitrary system to any desired frame of reference.

19-8 Analytical Self-Calibration Example

This section describes a self-calibration example using close-range terrestrial photography. The block included five photos taken with an off-the-shelf digital SLR camera. One of the photos used is shown in Fig. 19-7. There were 23 unique feature points used to generate 93 photo-coordinate pairs. One vertical control point and two three-dimensional control points were obtained using traditional field surveying techniques. The block is illustrated in Fig. 19-8. Photos 1 and 5 were about 20 m away from each other.

FIGURE 19-7 Photo 4 used in the adjustment.

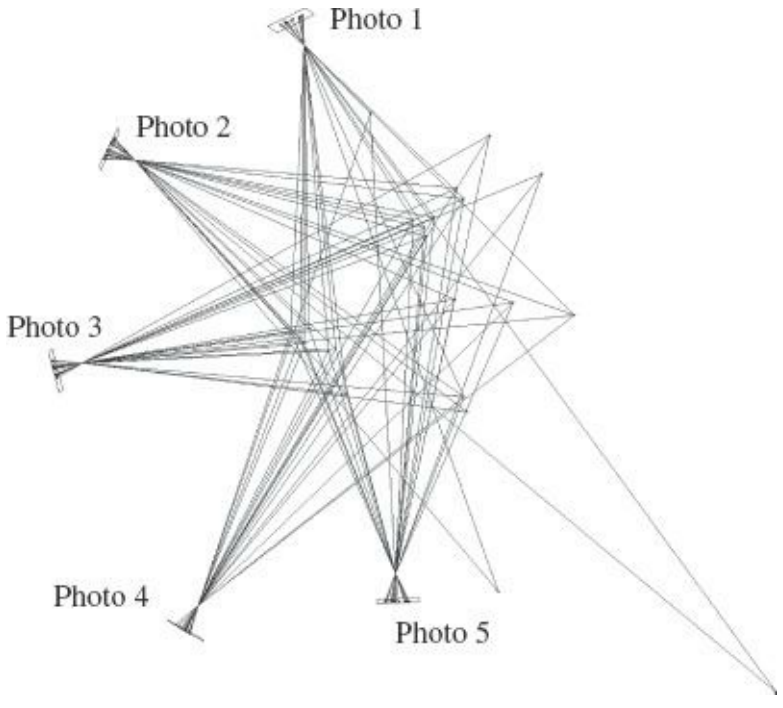


FIGURE 19-8 Graphical representation in plan view of the block of photos used in the example with 93 light rays representing the measured photo coordinate pairs.

Notice the convergent photography and varying depth of field leading to strong geometry for the adjustment. The focal length, f, was estimated using the manufacturer's specifications. In order to obtain pass points in millimeters—the units of f—the images were scaled and placed in a computer-drafting program, wherein the photo measurements were made, such that the ratio of focal length to format was nominally correct and the center of the photo was at the coordinate system origin. The rest of the camera calibration parameters were initialized at zero. Since the focal length was 7.1 mm and the equivalent focal length for 35 mm film is 28 mm, the equivalent format is 8.8 mm:

$$\frac{7.1 \text{ mm}}{28 \text{ mm}} 35 \text{ mm} = 8.8 \text{ mm}$$

Note that the above formula is purely based on manufacturer's specifications, and that the calibrated focal length is based on the photo measurements made on scaled photos. Thus the adjustment-resolved f, along with the other camera calibration parameters, is not the true physical focal length of the camera. However, this does not affect the adjusted object space coordinates of points since the calibration parameters are restricted to the image-space components of the collinearity equations. That is, the adjusted object space parameters are related only to the object space observations, which are the control points in this example. The initial approximations for the ground coordinates of pass points and exterior orientation parameters of the camera stations were found using manual methods described in Sec. 19-5.

On the first execution of the adjustment, the calibration parameters were constrained by giving the initial approximations very small standard deviations to help ensure convergence. The selected values for input and the output are shown in Table 19-1 and Table 19-2, respectively where σ_x and σ_y are the a priori standard deviations of the photo-coordinate measurements, and $\bar{\alpha}_X$, $\bar{\alpha}_Y$, and $\bar{\alpha}_Z$ are the average adjusted ground coordinate standard deviations.

Parameter	Value
f	Constrained
x_0	Constrained
y_0	Constrained
k_1	Constrained
k_2	Constrained
k_3	Constrained
p_1	Constrained
p_2	Constrained
σ_x, σ_y	0.005 mm

TABLE 19-1 Selected Input Parameters of the First Adjustment with All Camera Calibration Parameters Constrained

Output	Value
S_0	1.49
$\bar{\sigma}_X$	0.026 m
$\bar{\sigma}_Y$	0.028 m
$\bar{\sigma}_Z$	0.011 m

TABLE 19-2 Selected Output values of the First Adjustment with All Camera Calibration Parameters Constrained

Since the adjustment converged and there were no detectable blunders, the calibration parameters were loosened by effectively giving the initial approximations zero weight, allowing them to be adjusted. The selected input and output are shown in Tables 19-3 and 19-4, respectively.

Parameter	Value
f	Loosened
x_0	Loosened
y_0	Loosened
k_1	Loosened
k_2	Loosened
k_3	Loosened
p_1	Loosened
p^2	Loosened
σ_x, σ_y	0.005 mm

TABLE 19-3 Selected Input Parameters of the Second Adjustment with All Camera Calibration Parameters Loosened

Output	Value
S_0	0.41
$\bar{\sigma}_X$	0.016 m
$\bar{\sigma}_Y$	0.018 m
$\bar{\sigma}_Z$	0.005 m

TABLE 19-4 Selected Output of the Second Adjustment with All Camera Calibration Parameters Loosened

Notice in Table 19-4 that the adjusted object space coordinate standard deviations, $\bar{\alpha}_X$, $\bar{\alpha}_Y$, and $\bar{\alpha}_Z$, indicate more precise results than when the calibration parameters were constrained. Also, note that the standard error of unit weight, S_0, is lower than one indicating that the a priori standard deviations of the photo coordinate measurements were overly pessimistic. The t-statistics for the solved calibration parameters are shown in Table 19-5. The values revealed that y_0, p_1, and p_2, and were not

resolved such that their value was significantly different than their initial approximations at the 90 percent confidence level. This is due to insufficient redundancy and perhaps geometry of the photography to properly model these distortions.

Parameter	t-Statistic
x	2.34
x_0	2.22
y_0	1.08
k_1	2.23
k_2	2.09
k_3	2.43
p_1	1.27
p_2	0.53

TABLE 19-5 T-Statistics for the Second Adjustment with All Camera Calibration Parameters Loosened

Following an adjustment where y_0, p_1, and p_2 were removed from the adjustment by constraining their values to zero, it was found that k_2 was also not significant at the 90 percent confidence level with a t-statistic of 0.78. Its higher t-statistic in the second adjustment may have stemmed from correlation with, and compensation for, the three other insignificant parameters.

The final adjustment used the selected input parameters shown in Table 19-6. The insignificant calibration parameters were constrained and the a priori photo-coordinate standard deviations were lowered as shown in Table 19-7.

Parameter	Value
f	Loosened
x_0	Loosened
y_0	Constrained
k_1	Loosened
k_2	Constrained
k_3	Loosened
p_1	Constrained
p_2	Constrained
σ_x, σ_y	0.002 mm

TABLE 19-6 Selected Input Parameters of the Final Adjustment

Output	Value
S_0	1.07
$\bar{\sigma}_X$	0.008 m
$\bar{\sigma}_Y$	0.009 m
$\bar{\sigma}_Z$	0.004 m

TABLE 19-7 Selected Output of the Final Adjustment

A chi-squared test on S_0 indicated that the a priori photo coordinate standard deviations are appropriate since its value is not significantly different from one with 95 percent confidence. Notice that the adjusted ground coordinates of the points had horizontal standard deviations roughly half of those for the second adjustment. This shows the effect that overparameterization can have on the adjustment solution. Finally, the t-statistics in Table 19-8 confirm that the adjustable camera calibration parameters were significantly resolved.

Parameter	t-Statistic
f	2.88
x_0	3.87
k_1	22.00
k_3	3.85

TABLE 19-8 T-Statistics for the Second Adjustment with All Camera Calibration Parameters Loosened

Note that, unlike in aerial adjustments, the X and Y components of the ground coordinates are more weakly resolved than the Z component. This is a result of the different geometry of the terrestrial/close range configuration. The mostly horizontal direction of the light rays leads to weaker triangulation on the horizontal components relative to vertical. That is, errors in the image measurements lead to errors in the best estimate of the position of the intersection of rays more in the horizontal direction than the vertical.

19-9 Planning for Close-Range Photogrammetry

In some cases, close-range photogrammetric analyses are done with existing amateur photography. For example, a police officer may have taken photographs of the scene of a vehicle accident, and subsequent photogrammetric analysis may be needed to determine the positions of the vehicles, tire marks, and other evidence. Such photography may not be ideally exposed, well focused, or generally in an optimal geometric configuration for photogrammetric analysis. On the other hand, under controlled situations (e.g., industrial applications), preliminary planning can be performed so as to control such factors as type of camera, lighting, and camera orientation.

Three main considerations for pictorial quality are resolution, depth of field, and exposure.

Whether one is using digital or film cameras, resolution is important in that all points of interest must be clearly visible on the resulting image (see Sec. 3-14). A preliminary assessment should be made to ensure that the resolution is sufficient to adequately capture the smallest necessary details. Depth of field (see Sec. 2-3) is particularly important for ensuring proper focus. In close-range photogrammetry, since object depth is typically of significant size relative to the distance from the camera, a large *f*-stop setting may be required to ensure that the entire object is in focus. Proper exposure is necessary for the image points being measured to have sufficient contrast and definition. Particular attention should be paid to avoiding shadows on the object, especially when a flash is used for illumination. In some cases, special retro-reflective targets may be attached to points of interest prior to exposing the photographs. This allows the photographer to underexpose the background, with the targets remaining clear and bright.

There are also some physical constraints which must be considered in planning close-range photography. For example, object points must be visible from at least two (preferably more) photographs. A number of different camera locations may need to be considered in order to meet this constraint. Another constraint concerns the physical space around the object. In many applications, objects may be enclosed in confined spaces which makes effective determination of camera positions more difficult.

An important geometric consideration for close-range photography is the angular orientation of the camera exposure stations. Accuracy of the analytical solution depends, to a large extent, upon the angles of intersection between rays of light. The highest overall accuracy will be achieved when angles of intersection are near 90°. Figure 19-9*a* illustrates stereo photographic coverage of an object where the camera axes are parallel. In this figure, the parallactic angle f_1 to object point *A* is approximately 35°. In Fig. 19-9*b*, stereo coverage of the same object is obtained from convergent photos. In this figure, the corresponding parallactic angle f_2 is approximately 95°. Since f_2 is closer to 90° than f_1, the overall accuracy of the computed coordinates of point *A* will be higher in the configuration of Fig. 19-9*b*. Notice also that the stereoscopic coverage in Fig. 19-9*a* is only approximately 50 percent of the field of view, while, the stereoscopic coverage of Fig. 19-9*b* is 100 percent of the field of view. This enables the full format of the camera to be used, resulting in greater efficiency and higher effective image resolution.

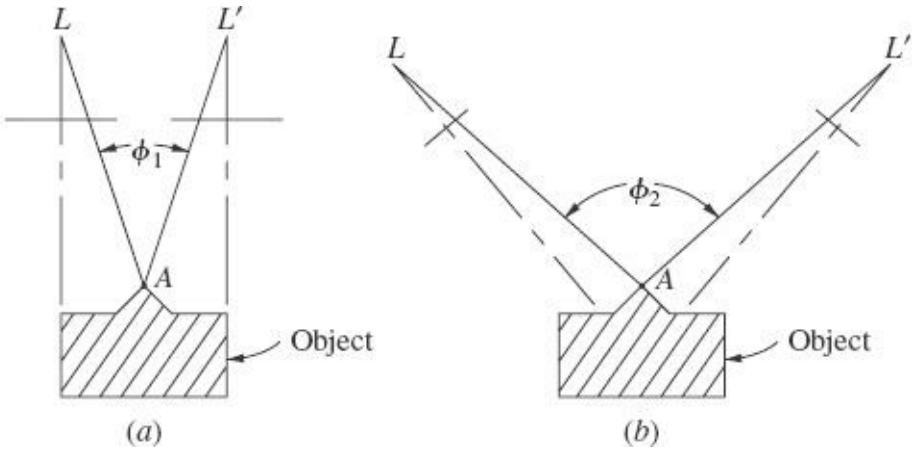


FIGURE 19-9 (*a*) Close-range stereo coverage of an object with parallel camera axes. (*b*) Close-range stereo coverage with convergent photography.

For the highest level of accuracy in close-range photogrammetry, a fully analytical solution is preferred. By applying the principles of analytical photogrammetry as described in Chap. 11 and Sec.

17-6, precisely measured photo coordinates of images can be used to directly compute *X*, *Y*, and *Z* coordinates in object space. The foundation of the analytical solution is the collinearity condition which gives rise to the collinearity equations [see Eqs. (11-1) and (11-2)]. These equations can be directly applied to terrestrial as well as aerial photographs.

In the preferred analytical method, the self-calibration approach described in Sec. 19-4 is used. This gives a calibration of the camera under the actual conditions (temperature, humidity, etc.) which existed when the photographs were taken. Certain geometric requirements must be met in order to effectively perform analytical self-calibration. First, numerous redundant photographs from multiple locations are required, with sufficient *roll diversity*. Roll diversity is a condition in which the photographs have angular attitudes that differ greatly from each other. Another requirement is that many well-distributed image points be measured over the entire format. This is important for accurate determination of lens distortion parameters.

Accurate measurement of photo coordinates is necessary to ensure accurate results from the analytical solution. High-precision comparators are generally used for film-based photographs. Digital camera systems, on the other hand, rely upon image-matching techniques (see Sec. 15-8) to obtain accurate photo coordinates. In any case, it is essential to properly identify object points as they appear on the different photos. Mislabeled points will result in an inaccurate analytical solution or, in some cases, will cause the solution to fail completely.

References

American Society of Photogrammetry: *Manual of Photogrammetry*, 5th ed., Bethesda, Md., 2004.

———: *Handbook of Non-Topographic Photogrammetry*, 2d ed., Bethesda, MD, 1989.

Brown, D. C.: "New Developments in Photogeodesy," *Photogrammetric Engineering and Remote Sensing*, vol. 60, no. 7, 1994, p. 877.

———: "Close-Range Camera Calibration," *Photogrammetric Engineering*, vol. 37, no. 8, 1971, p. 855.

Fischler, M.A., R.C. Bolles: "Random Sampling Consensus—a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Communications of the ACM*, vol. 24, no.6. 1981. p. 381.

Fraser, C. S., H. Hanley, S. Cronk:"Close-range Photogrammetry for Accident Reconstruction," Proceedings of Optical 3D Measurements VII. Vienna, Austria. 2005. p. 115.

Fraser, C.S.: "Some Thoughts on the Emergence of Digital Close Range Photogrammetry," *Photogrammetric Record*, vol. 16 no. 91, 1998. p. 37.

C. S. Fraser, "Digital Camera Self-calibration," *ISPRS Journal of Photogrammetry and Remote Sensing*. vol. 52, no. 4, 1997, p. 149.

Fraser, C. S.: "Photogrammetric Measurement to One Part in a Million," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 3, 1992, p. 305.

———: "Microwave Antenna Measurement," *Photogrammetric Engineering and Remote Sensing*, vol. 52, no. 10, 1986, p. 1627.

———: "Photogrammetric Measurement of Thermal Deformation of a Large Process Compressor," *Photogrammetric Engineering and Remote Sensing*, vol. 51, no. 10, 1985, p. 1569.

———: "Network Design Considerations for Non-Topographic Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 8, 1984, p. 1115.

———: "Optimization of Precision in Close-Range Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 48, no. 4, 1982, p. 561.

Fryer, J. G., and D. C. Brown: "Lens Distortion for Close-Range Photogrammetry," *Photogrammetric*

Engineering and Remote Sensing, vol. 52, no. 1, 1986, p. 51.
Ghosh, S. K.: "Photogrammetry for Police Use: Experience in Japan," *Photogrammetric Engineering and Remote Sensing*, vol. 46, no. 3, 1980, p. 329.
Karras, G. E., and E. Petsa: "DEM Matching and Detection of Deformation in Close-Range Photogrammetry without Control," *Photogrammetric Engineering and Remote Sensing*, vol. 59, no. 9, 1993, p. 1419.
Kenefick, J. F., M. S. Gyer, and B. F. Harp: "Analytical Self-Calibration," *Photogrammetric Engineering*, vol. 38, no. 11, 1972, p. 1117.
Lichti, D. D., and M. A. Chapman: "Constrained FEM Self-Calibration," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 9, 1997, p. 1111.
Okamoto, A.: "The Model Construction Problem Using the Collinearity Condition," *Photogrammetric Engineering and Remote Sensing*, vol. 50, no. 6, 1984, p. 705.
Remondino, F., C. Fraser: "Digital Camera Calibration Methods: Considerations and Comparisons." Proceedings of ISPRS Commission V Symposium, Dresden, Germany, 2006.
Roth, D. G.,:"Automatic Correspondences for Photogrammetric Model Building," Proceedings of the XXth ISPRS Congress, Istanbul, Turkey. 2004.
Stewenius, H., C. Engels, and D. Nister, "Recent Developments on Direct Relative Orientation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 4, 2006. p. 284.
Wong, K. W.: "Machine Vision, Robot Vision, Computer Vision, and Close-Range Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 8, 1992, p. 1197.
———: "Mathematical Formulation and Digital Analysis in Close-Range Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 11, 1975, p. 1355.

Problems

19-1. Discuss some of the uses of terrestrial or close-range photogrammetry.

19-2. Describe the differences between metric and nonmetric terrestrial cameras.

19-3. Describe the pros and cons of calibrating a camera using mission data versus calibrating a camera in a laboratory environment.

19-4. Explain how calibration parameters can be constrained in a self-calibrating bundle adjustment.

19-5. Explain why the tilt-swing-azimuth system is useful for finding initial approximations of the attitude angles for terrestrial photography.

19-6. Describe why starting with two images and sequentially adding photos to an adjustment is a good strategy for performing a close range adjustment with self-calibration.

19-7. Why is it advisable to constrain p1 and p2 when x0 and y0 are loose in adjustments with little redundancy.

19-8. Discuss four basic approaches in establishing control for terrestrial or close-range photogrammetry.

19-9. Name and discuss three considerations which affect pictorial quality that should be considered in planning close-range photography.

19-10. Discuss how accuracy can be improved through the use on convergent photos versus stereo photographs with parallel camera axes.

19-11. Describe two geometric requirements for photography when analytical self-calibration is used for computing object coordinates in close-range photogrammetry.

19-12. Explain how incorrectly-scaled image coordinates affect the solved focal length and object space coordinates of control points in a bundle adjustment with self-calibration.

CHAPTER 20
Photogrammetric Applications in GIS

20-1 Introduction

As noted in earlier sections of this book, photogrammetry and remote sensing play extremely important roles in the development and implementation of geographic information systems. In one very useful application, aerial images are frequently employed by GIS operators as background frames of reference for performing spatial analyses. The images may be obtained from either aerial cameras or satellite systems, and in general they will have undergone some form of georeferencing such as conversion to digital mosaics (see Chap. 9) or orthophotos (see Chap. 13) prior to being used in this manner. But perhaps the most important contribution to geographic information systems made by photogrammetry and remote sensing is their use in directly generating spatially accurate feature information for databases.

Information for specific GIS databases can be compiled directly from stereomodels created from aerial photographs by photogrammetric restitution instruments such as analytical plotters or softcopy plotters (see Chaps. 12 and 13). As examples, layers of information such as roads, hydrography, land use, and many others can be coded and digitized directly from stereomodels, and entered into GIS databases. Digital line graphs (DLGs) and digital elevation models (DEMs) are two additional products that are frequently compiled directly from stereomodels. The former gives planimetric positions of objects, while the latter provides elevation data. Both are essential for many GIS applications. The digital orthophoto (see Chap. 13) is another photogrammetric product that has become indispensable in GIS work. Digital orthophotos can be automatically and very economically compiled. They are especially valuable because they are in image (raster) form, and GIS operators can analyze them visually. In addition they are planimetrically accurate, and thus two-dimensional features can be digitized directly from them. Because of these characteristics, they are extremely convenient for use as background reference frameworks for GIS applications. Roads, railroads, wetland boundaries, and other planimetric features that appear in raster form on a digital orthophoto, for example, can readily be converted to vector form so that they are amenable to various GIS spatial analyses.

Layers of information for geographic information systems are often compiled by simultaneously analyzing photogrammetric products in conjunction with other documents. As an example, a land ownership layer may be conveniently developed by making a visual analysis of digital orthophotos while simultaneously reading and interpreting parcel descriptions, dimensions, and other information given in deeds, plat books, and tax documents. Generating GIS database information by photogrammetry is almost always faster and more cost-effective than doing so by any other means, and the process enables high orders of accuracy to be achieved.

In sections that follow, a variety of applications are presented in which geographic information systems were used as a means of solving problems. The contributions made by photogrammetry and remote sensing in these applications are described.

20-2 Land and Property Management

The local government of St. Johns County, Florida, encompasses 609 mi^2 and serves a population of

approximately 200,000. Services for its residents and visitors require an inventory of public lands, buildings, streets, and other resources which require a variety of applications of GIS for management and operation. Organized records of buildings along with measurements and spatial analyses support operations and maintenance. Aerial image products, particularly orthophotos, derived from photogrammetric methods support the layer of all county-owned parcels that include data regarding size, current use, future use, and legal requirements. In addition, stereo-plotting was used to accurately digitize street elements (pavement, curb and gutter, etc.), stormwater retention ponds, building footprints, and various elevation data which are used to directly manage county land, buildings, and facilities, as well as analyze and enhance the use of these features.

The GIS is the primary site development and analytical tool for informed land management. County personnel need to analyze large areas of land to determine potential sites for new facilities such as roads, stormwater drainage systems, utilities, and communication towers. For site selections, staff must first determine if any existing county property would be suitable. Data organized in a GIS readily enable queries to be made to eliminate many inappropriate parcels based on design constraints and legal considerations. GIS spatial analysis facilitates overlaying imagery, parcels of land, storm surge prediction areas, flood zones, National Wetland Inventory estimates, land cover, elevation, zoning, address points, and utility lines along with the remaining county-owned parcels. Search buffers and road network routing support efficient site selection.

These site analysis tools and data supported the design of ten sites for the county intergovernmental radio system. Telecommunications engineering and modeling determined approximate locations for towers. Using GIS layers for research and spatial analysis helped ensure a full review of land was performed and the best available sites were selected so that this important system could be developed. (See Fig. 20-1.)

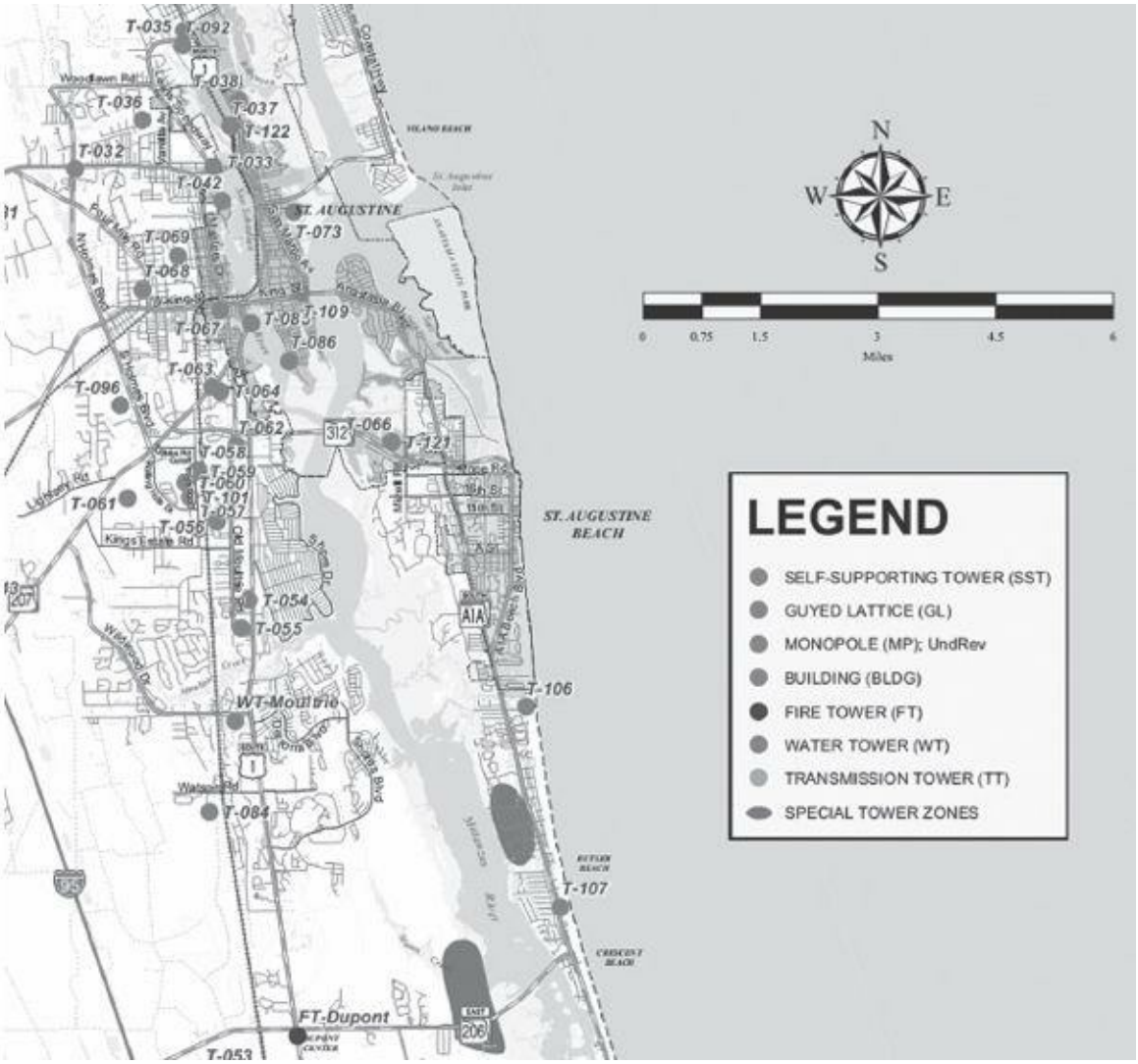


FIGURE 20-1 Telecommunication tower sites for a portion of St. Johns County, Florida.

20-3 Floodplain Rating

The Federal Emergency Management Agency (FEMA) works with communities across the United States to help homeowners, contractors, architects, local codes officials, and engineers build smarter and safer homes. FEMA also manages the National Flood Insurance Program (NFIP) which provides homeowners, renters, and business owners a way to protect themselves financially from catastrophic flood losses. The NFIP's community rating system (CRS) is a voluntary incentive program that recognizes and encourages community floodplain management activities that exceed the minimum NFIP requirements. As a result, flood insurance premium rates are discounted to reflect the reduced flood risk resulting from various community actions.

St. Johns County, Florida, participates in the CRS program to help reduce flood damage losses, provide proper regulation for construction, increase public awareness of flood zone locations and impacts, and prepare for flooding situations. By participating in the CRS program, damage from flooding can be minimized and property owners' flood insurance rates reduced to reflect the lower damage impacts. The CRS contains extensive guidelines and recognition for improvements within the categories defined in the NFIP. Photogrammetry, remote sensing, and LiDAR products incorporated in the county GIS are used to more accurately delineate floodplains and identify potential improvements. (See Fig. 20-2.)

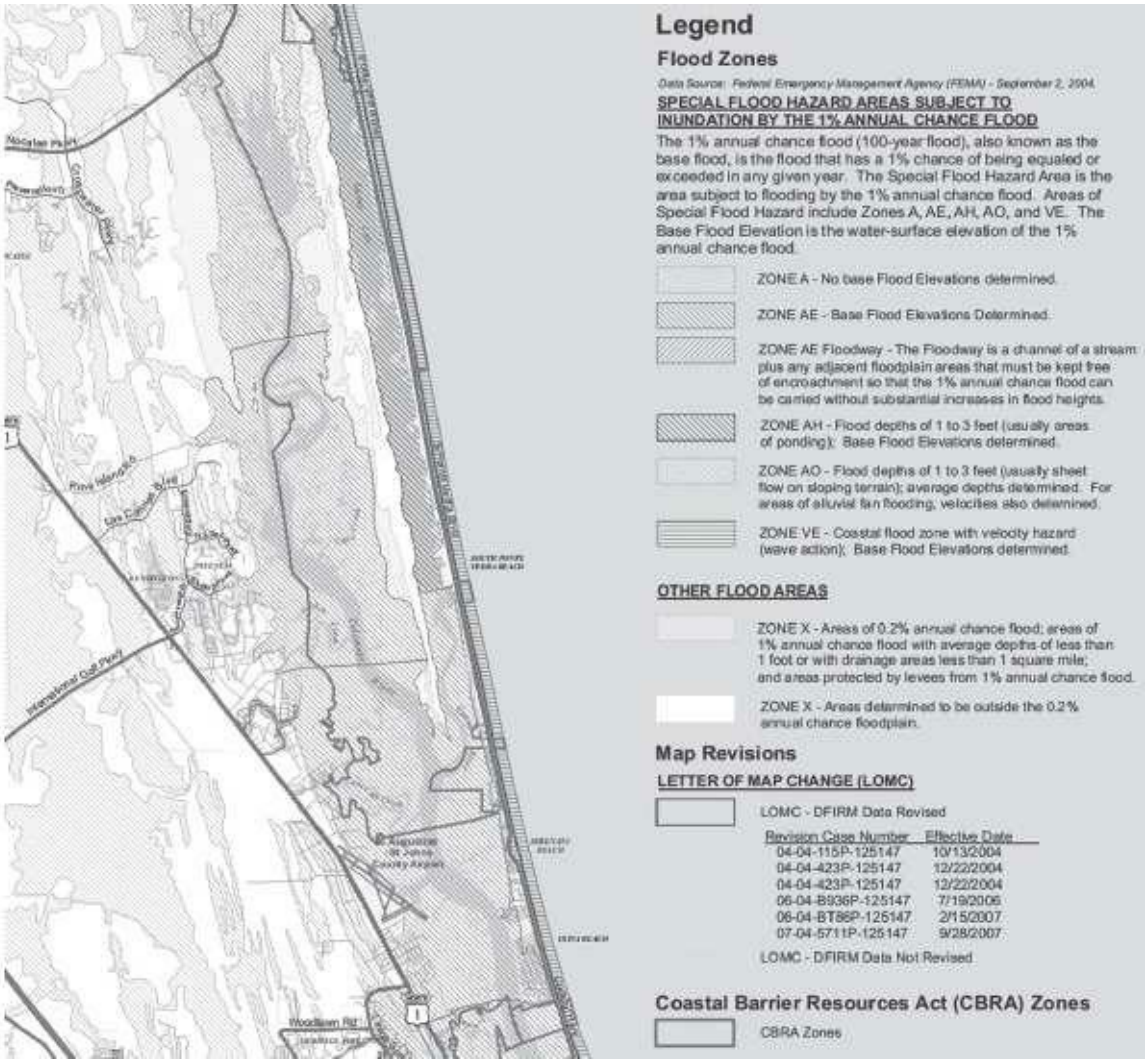


FIGURE 20-2 Floodplain ratings for areas within St. Johns County, Florida.

In addition, the GIS can be used to efficiently determine how much open space exists within the regulatory floodplain areas. Increasing open space within the floodplains helps reduce the potential for damage and increases available space for flood waters. In addition to identifying the floodplain areas and open space, the GIS incorporates private land ownership, public lands, and land uses in order to maintain an effective CRS. Water bodies were delineated and land uses updated from orthophotography, which was used as a base map for the entire project. Structures including roads and buildings were also reviewed and removed, and designation as conservation lands increased the value of the preserved open space. Knowledge of the floodplains, open space, preserved lands, and percentage of floodplain lands covered by these open spaces has helped St. Johns County protect property, reduce insurance costs, and provide safer conditions for homes and businesses.

20-4 Water Quality Management

Geographic information systems have been successfully applied to the solution of many problems in water quality management. A project of this type selected for discussion in this text involves a GIS that was developed for the Sugar River watershed in south-central Wisconsin (see Fig. 20-3). The objectives of the GIS were to provide a cost-effective means for assisting in preventing soil erosion and to improve and preserve the water quality in the Sugar River and its tributaries, as well as in Belle Lake which exists at the lower end of the watershed. The GIS was developed by the Dane County Land

Conservation Department, with funding from the Wisconsin Department of Natural Resources, the U.S. Environmental Protection Agency, and Dane County.

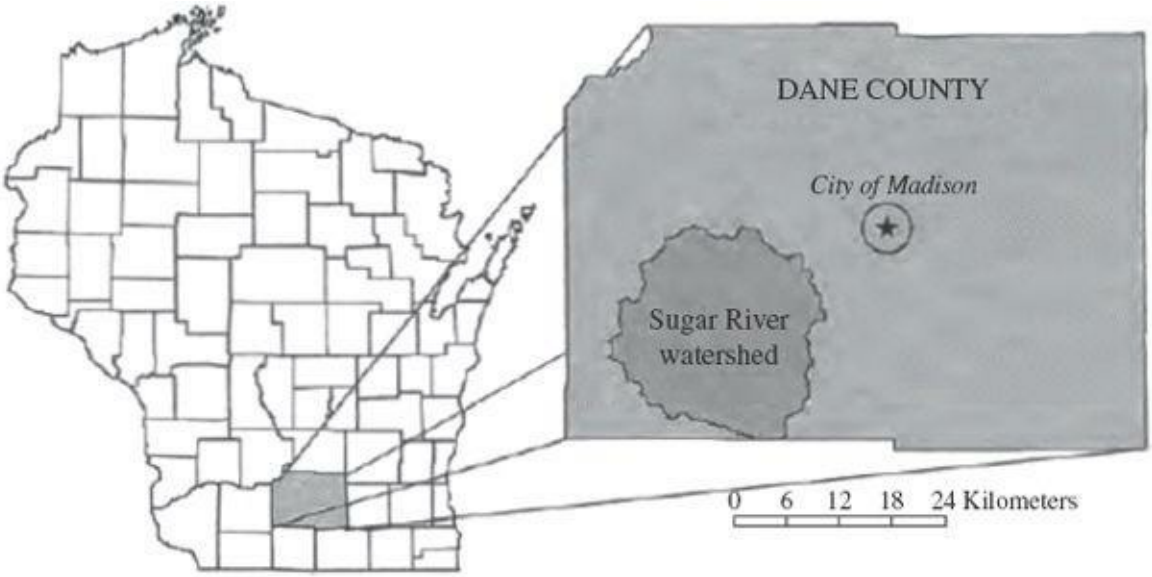


FIGURE 20-3 Location of Sugar River water quality management GIS project. (*Courtesy Dane County Land Conservation Department, M. S. Robinson and A. Roa-Espinosa*.)

In the Sugar River watershed, excessive amounts of sediment and phosphorus, generated by both agricultural and urban land uses, enter the surface water. These conditions are detrimental to the aquatic and wildlife habitat within the basin, and they also impact negatively upon the aesthetic qualities and recreational use of the water. To address this problem, a GIS was developed to provide the information necessary for a computer modeling program to identify areas of high sediment and phosphorus delivery within the watershed. Improved management practices could then be implemented in these areas.

The major layers of information developed for the GIS database included topography, hydrography, soils, and land use/land cover. Photogrammetry played a major role in the development of each of these layers. The elevation layer (see Fig. 20-4) was based upon a digital elevation model that was produced photogrammetrically. The DEM, which consisted of a 10-m grid with accuracies to within ±0.3 m, provided slope information that was critical in the computer modeling. The hydrography layer, which consisted of the rivers and streams within the watershed, was digitized from orthophotos. By combining the elevation and hydrography layers, small individual hydrologic drainage areas that were needed for the computer modeling were identified. There were approximately 500 individual areas in the watershed, as shown in Fig. 20-5, and they averaged about 1 km^2 in size. The soils layer was digitized from NHAP aerial photos.

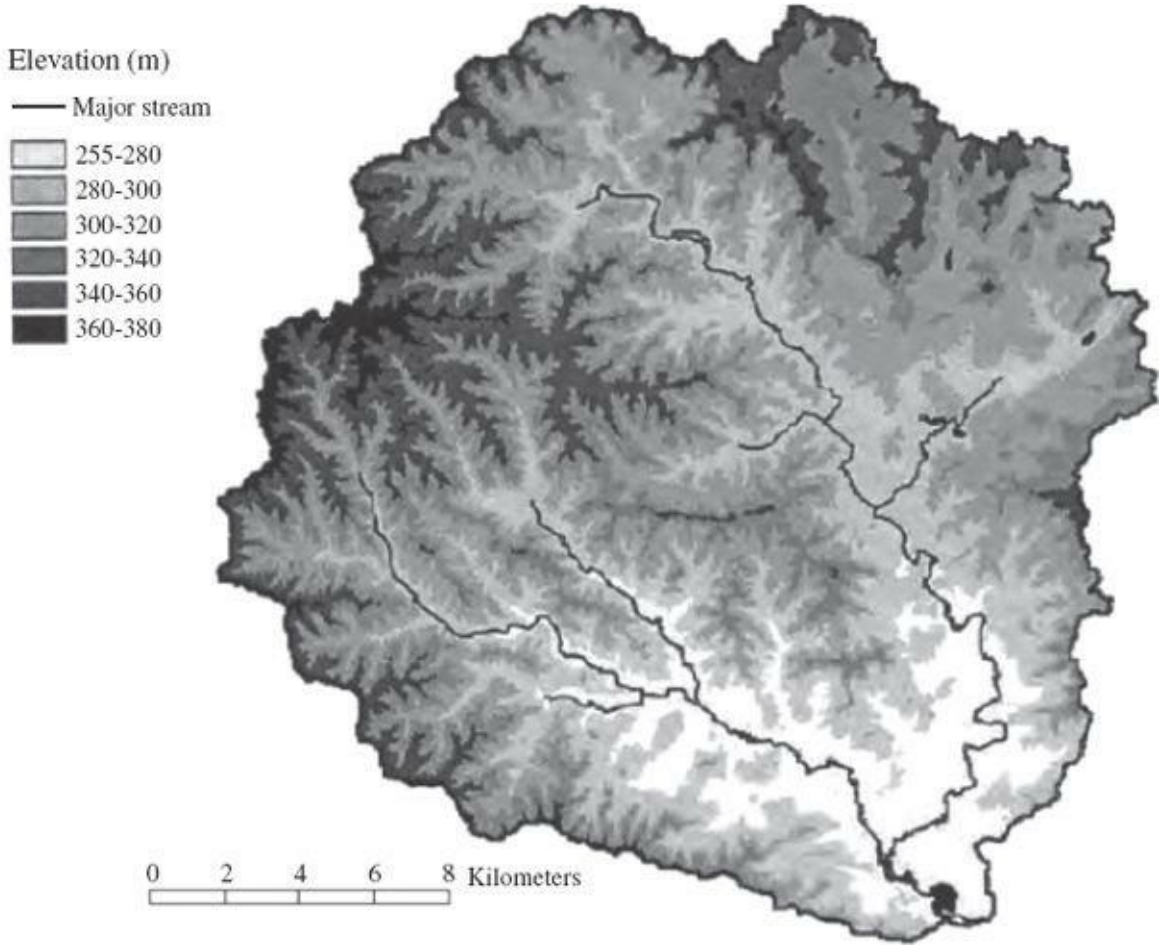


FIGURE 20-4 Elevation model prepared from DEM of Sugar River watershed. (*Courtesy Dane County Land Conservation Department, M. S. Robinson and A. Roa-Espinosa*.)

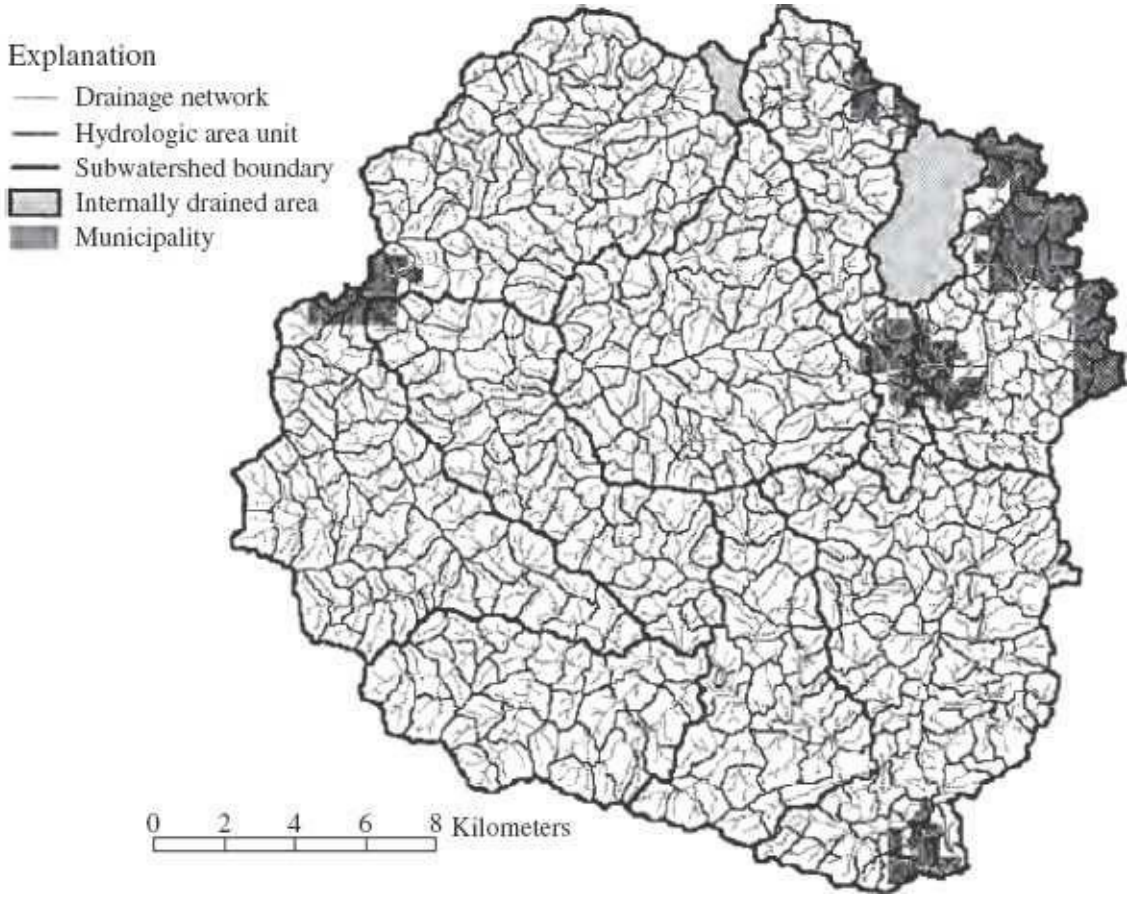


FIGURE 20-5 Hydrologic area units within the Sugar River watershed obtained by combining

elevation and hydrography layers. (*Courtesy Dane County Land Conservation Department, M. S. Robinson and A. Roa-Espinosa*.)

The land use/land cover layer consisted of the combination of many individual layers. It included categories of cropland, grassland, pasture, woodland, wetland, water, roads, residential areas, industrial areas, and commercial areas (see Fig. 20-6). Although other sources such as zoning maps and parcel maps were used to compile this information, most of these data were obtained from orthophotos. Individual farm tracts were included within this layer, and these tracts were further subdivided into separate fields. This was important because their differing crop rotations and land uses have an important bearing on their rates of sediment and phosphorus delivery.

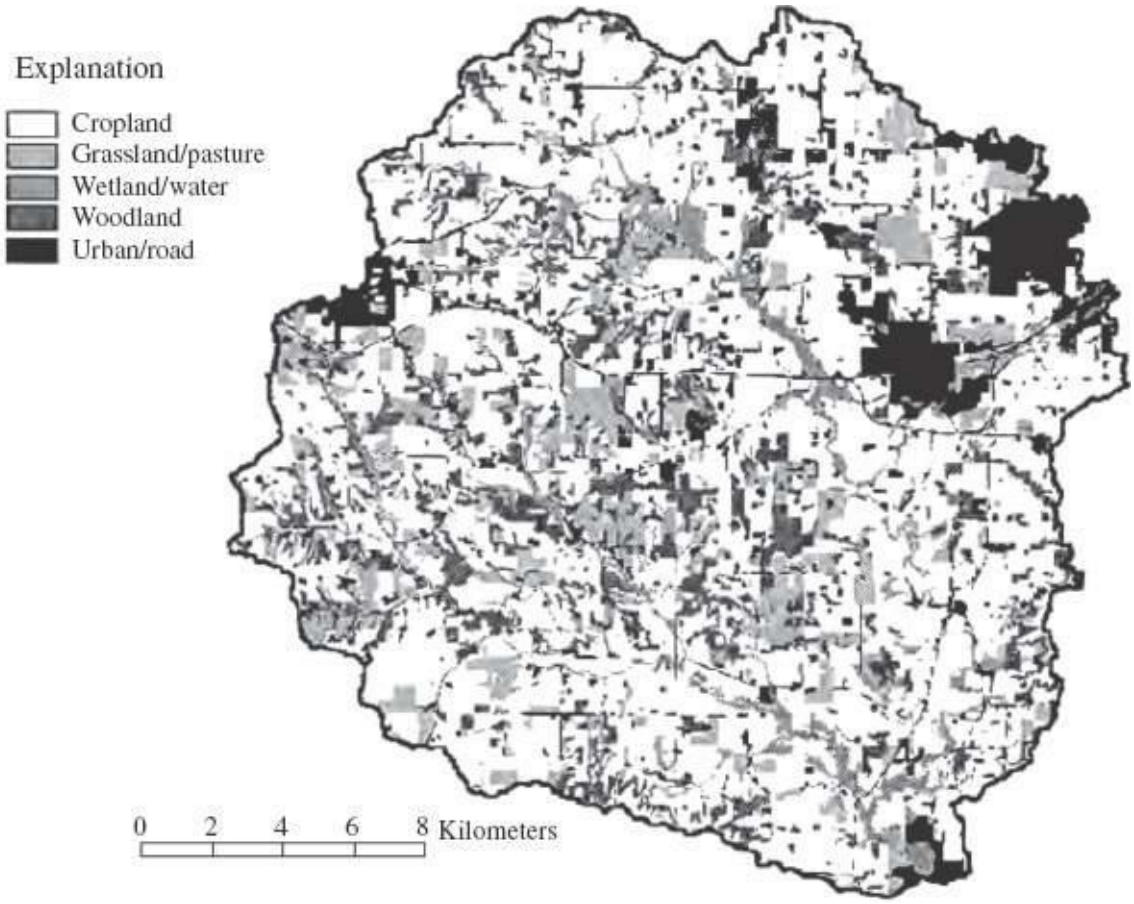


FIGURE 20-6 Land use/land cover layer of the Sugar River GIS database. (*Courtesy Dane County Land Conservation Department, M. S. Robinson and A. Roa-Espinosa*.)

The previous information was used by the computer program to model the sediment and phosphorus delivery rates across the watershed. Figure 20-7 is a graphic which shows the sediment yield of each of the hydrologic units in the watershed. A similar graphic was developed for phosphorus yields. Areas with high yields could be examined in the field and targeted for improved field and management practices. These practices could include such measures as practicing conservation tillage, using soil binders, establishing buffer strips, constructing terraces, or installing water and sediment control basins.

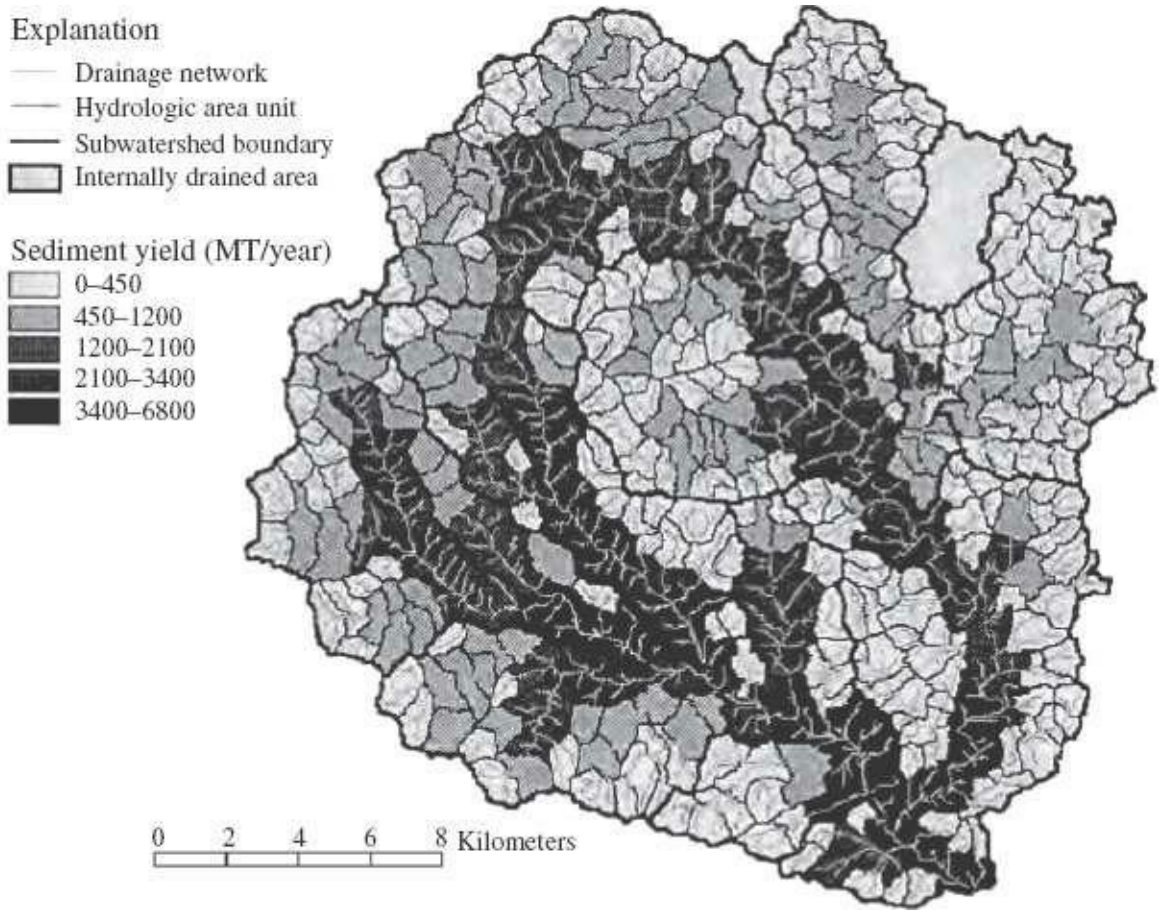


FIGURE 20-7 Sediment yield within the Sugar River watershed obtained through GIS analysis. (*Courtesy Dane County Land Conservation Department, M. S. Robinson and A. Roa-Espinosa*.)

20-5 Wildlife Management

An example illustrating a GIS application in wildlife management involves a portion of the Lower Virgin River in the southwestern United States. The study area, shown in Fig. 20-8, includes the entire riparian corridor of the river from the Virgin River Gorge near Little-field in northwest Arizona, to the Lake Mead National Recreation Area in southeast Nevada. This involves a total of 12,349 acres of riparian vegetation and open water/river channel over a length of approximately 52 mi of river. The study was conducted by the Southern Nevada Water Authority. Its objectives were to delineate the vegetation communities and other habitat types (open water, etc.) within the riparian corridor of the river, determine how the different vegetation species are used by wildlife, and predict how the distribution and composition of vegetation and wildlife change with varying flow levels in the river. Evaluations of the status of the federally endangered southwestern willow flycatcher and the federally threatened desert tortoise were matters of special interest in this project.

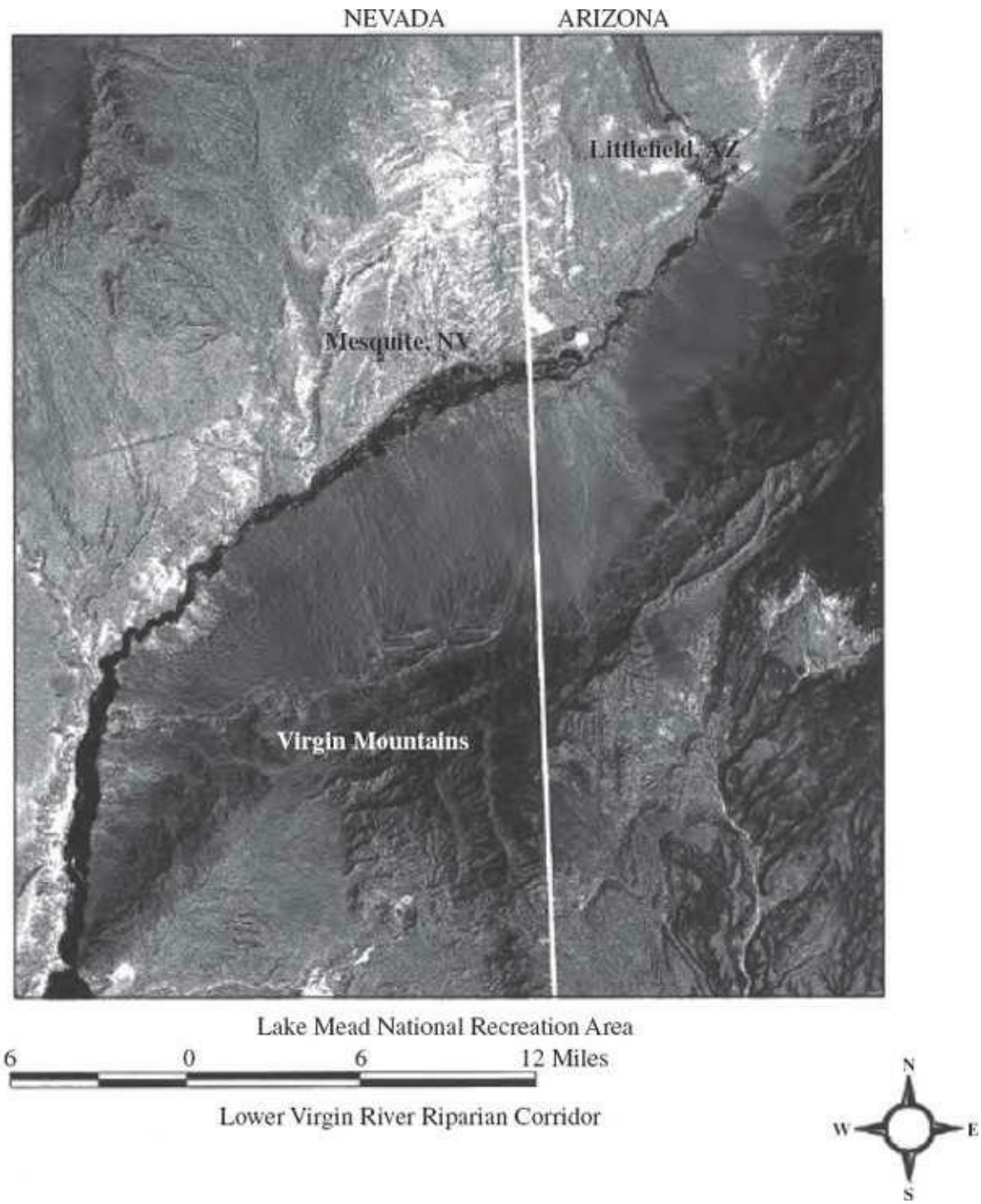


FIGURE 20-8 Study area of Virgin River wildlife management GIS project. (*Courtesy Southern Nevada Water Authority*.)

To achieve the project's goals, a GIS was developed which covered the area. The base map, or frame of reference for performing GIS analyses, was prepared from color infrared aerial photos. These photos, taken at a scale of 1:24,000, were scanned, rectified, and joined digitally to form a mosaic of the study area. A total of 38 photos were involved in the mosaic construction. Control for the digital mosaic process was obtained from USGS digital line graphs.

In this study, one of the critical layers of information compiled was vegetation. To assist in preparing this layer, portions of the base map in selected areas were enlarged to 1:6000 scale. This increased the visual resolution which facilitated their use for fieldwork. These large-scale photomaps were used during field surveys to identify and delineate the locations of vegetation classifications in several key areas of the river corridor. Based upon these "ground truth" samples, the remaining images in the corridor were interpreted and classified, using the heads-up digitizing process (see Sec.

9-4). A hard copy of this preliminary vegetation map was then taken into the field, and the delineations of vegetation types were verified and modified as necessary. GPS was also used in some cases to check visual analyses and to locate certain features that were not visible on the base map, such as special habitat areas and nesting grounds. These modifications were then incorporated into the vegetation layer of the database. Next the resulting vegetation map was utilized to select representative areas for evaluating wildlife and their use of the resources within the riparian corridor. A total of 11 vegetation classifications were included on the map. Figure 20-9 shows the portion of the vegetation map at the location where the Virgin River corridor enters Lake Mead. Note that only five classifications existed within this section of the corridor.

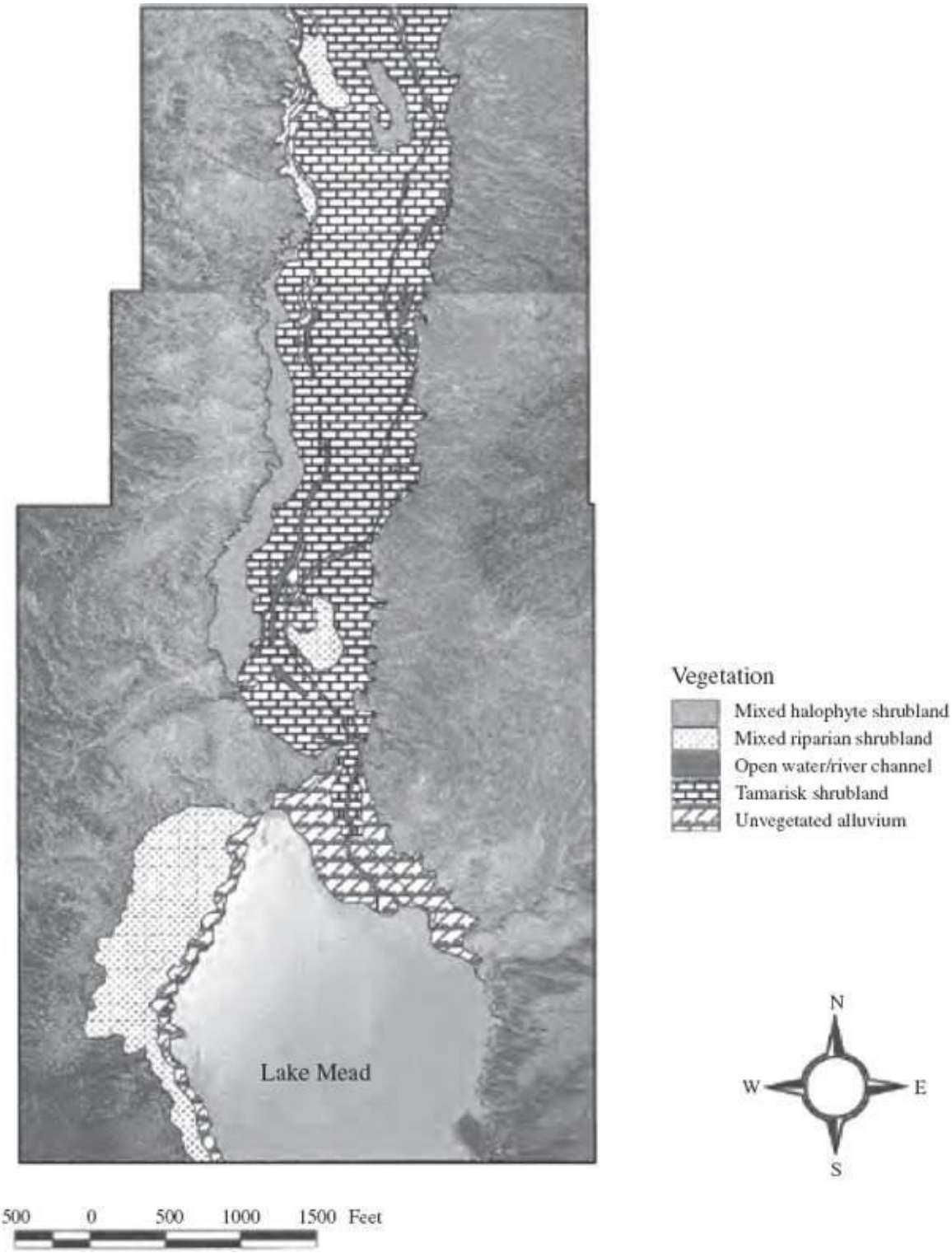


FIGURE 20-9 Vegetation layer of Virgin River wildlife management GIS project. (*Courtesy Southern*

Nevada Water Authority.)

The vegetation/habitat layer was used as a guide in selecting areas for conducting wildlife surveys in the field. The areas were chosen so that each of the vegetation classifications and habitats were represented within the survey. In addition to vegetation types, the topography, floodplain characteristics, and stream gradient were also considered in selecting the wildlife survey areas. Accessibility was a consideration as well. Sampling transects which included small mammal live traps, pitfall traps, and bird observation points were placed at selected stations within the survey areas to sample the wildlife populations. GPS was used to control the placement and mapping of the transect lines for data capture within the GIS model. To date, a total of 170 different species of birds, 34 different species of mammals, 13 species of reptiles, and 4 species of amphibians have been found to utilize the Lower Virgin River riparian corridor.

This project has been meeting all its objectives. A great deal of valuable data have been gathered to further existing knowledge of the wildlife in this southwest habitat, and to aid in understanding the dependence these wild creatures have upon the different vegetation communities and other habitat types that exist in the river corridor. The study of this 52-mi-long riparian corridor would have been a daunting task if performed by ground methods alone. However, as discussed, photogrammetry has contributed significantly in the conduct of this study. Not only has photogrammetry enabled the work to be completed more rapidly and efficiently, but also it has permitted greater amounts of data to be collected, analyzed, and managed, thereby improving the overall end results. This method of study has the added advantage that the information collected can conveniently be shared with other agencies having overlapping or common interests and goals.

20-6 Environmental Monitoring

The Southern Nevada Water Authority (SNWA) is responsible for supplying water to nearly 2 million people in the Las Vegas region. Ninety percent of the water serving the region comes from a single source, the Colorado River. Since 1999, the river has undergone extreme drought, with the quantity of water diminishing by nearly 50%. The SNWA has been looking for other sources to help supplement existing water supplies, and groundwater in valleys north of Las Vegas has been considered as a potential option.

The SNWA looked to meet the water needs of nearly 2 million people while maintaining natural habitat and minimizing negative impacts to the environment. This was done through environmental monitoring for the large area of interest to SNWA. Analysis of large areas by performing regular field surveys is generally not feasible, as it would take an enormous effort by numerous staff members to effectively monitor such areas. However, large-scale monitoring has become easier over the past few decades, as the advent of commercial airborne imagery allows companies to view and analyze thousands of square miles in a much shorter timeframe. Satellite imagery can be periodically acquired over the same areas, allowing for analysis to detect changes that have occurred over time. Spatially accurate multispectral imagery, once the exclusive domain of satellites, now can also be routinely acquired from airborne sensors. This allows for better resolution of imagery and reduces the detrimental atmospheric effects inherent in satellite imagery.

As part of the monitoring efforts in basins where groundwater could be pumped, SNWA has collected baseline data over areas of eastern Nevada. Among the data collected were both multispectral imagery and LiDAR elevations. The image resolution (ground sample dimension) selected for this project was 6 in (15 cm) and the LiDAR data were collected at a nominal spacing of

1.1 m. Together, these data helped establish a high-quality baseline for areas of interest to SNWA. The collection of the data was a large task, as nearly 3000 mi^2 of LiDAR and imagery were acquired over a 4-week period wherein the longest length was about 125 mi and the greatest width 80 mi. The area of interest is very sparsely populated, has minimal transportation coverage, and ranges in elevation from 5500 to 13,000 ft.

The initial spatial control of aerial-based imagery and LiDAR data was positioned by airborne GPS (AGPS) systems deploying strategically placed ground receivers that coordinated with the flight team and flight plan as a function of time and schedule. In addition to in-flight AGPS, staff members were deployed to collect spectrometer readings of plant and ground features to help identify ground cover, assist with color-balancing of imagery, and aid in plant classification. Once acquisition of imagery and LiDAR was completed, statically collected GPS control and checkpoint targets were compared against AGPS's kinematic positions. Out of the nearly 100 checkpoints, 99% were within the ASPRS (American Society for Photogrammetry and Remote Sensing) horizontal and vertical accuracy requirements, and most were within a few centimeters of static GPS-measured coordinates.

After validating the spatial accuracy of the dataset, the checkpoint positions were utilized as supplemental control and the processing was redone to produce the final products. It is also important to understand that the use of ground and AGPS provided a reliability factor in the event that the AGPS system failed during flight, thereby ensuring the expensive field deployment would be successful and would produce accurate results. Processing of both datasets took several months, but was ultimately completed and accepted by SNWA.

These datasets have been instrumental in different aspects of the monitoring program, including GIS mapping and analysis for hydrology, biology, and ranch management. Spring locations, property boundaries, plant identification, soil analysis, slope analysis, animal grazing allotments, and determining the volume of small drainage basins are just some of the items that have been determined or benefitted from the more accurate datasets that were collected. Prior to this acquisition, the best imagery available in the area was 1-m resolution, a good product for a large area but it lacked detail that was needed for some of the analyses. Equally important was the elevation grid that was generated from the LiDAR data. This grid greatly improved upon the existing 30-m grid created in the early 1990s. Figure 20-10 depicts the imagery produced for the project area.

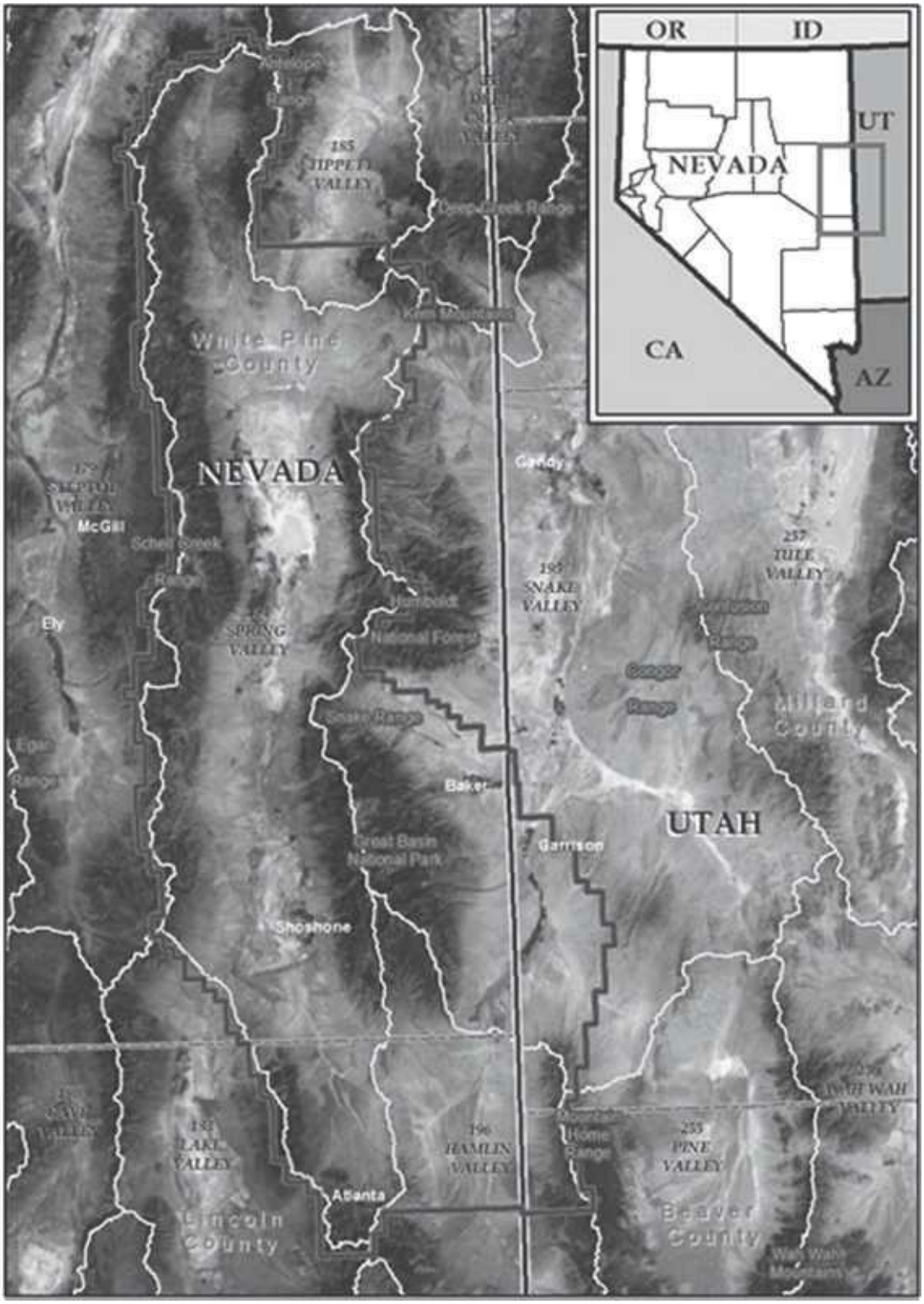


FIGURE 20-10 Groundwater resource area in eastern Nevada being monitored for environmental effects. (*Courtesy Southern Nevada Water Authority*.)

20-7 Wetland Analysis

It has been well established that wetlands provide crucial benefits to the natural environment. Besides providing habitat for fish and other wildlife, they act as “filters” to impede stormwater runoff contaminants from reaching the groundwater and surface waters. In St. Johns County, Florida, as well as most areas of the country, wetland delineation and analysis is therefore important for facilitating informed and responsible land management. Land uses and management practices must not cause detrimental impacts to these wetlands. For example, stormwater runoff from parking lots often

contains toxic chemicals that can be very harmful to plants and wildlife. Runoff from landscaped areas often contains herbicides that can also have similar negative impacts. Delineation of wetlands and analysis of hydrological characteristics are considered essential for the preservation of these important areas.

Detailed wetland analysis involves substantial knowledge and experience. A GIS can be a useful tool to aid the wetland scientist in finding general locations for analysis. When delineating wetlands on a parcel of county-owned land, the starting point is always GIS. Some of the key layers in refining the analysis and directing ground analysis include: land cover, national wetlands inventory (NWI), soils, true color imagery, color infrared imagery, and LiDAR elevation data. Starting with the land cover, soils, and NWI layers, the nominal habitat type and potential hydric indicators (particularly hydric soils or hydrophytic vegetation) are checked to indicate the possible presence of wetlands. Subsequently, the imagery layers provide additional clues allowing the wetland scientist to recognize certain image characteristics as indicative of wetlands. Elevation data often refines the search area for wetland delineation when combined with aerial imagery. The combination of these indicators can provide a nominal indicator of the wetland boundary. With a preliminary GIS review of the parcel, time in the field is directed to the areas needing further scrutiny and the field process of wetland delineation is much more efficient and accurate. By measuring the spatial location of the boundary in the field, a more accurate representation can be input into the GIS database, which will improve the accuracy of the information. Figure 20-11 shows a portion of the wetlands layer in the St. Johns County GIS.

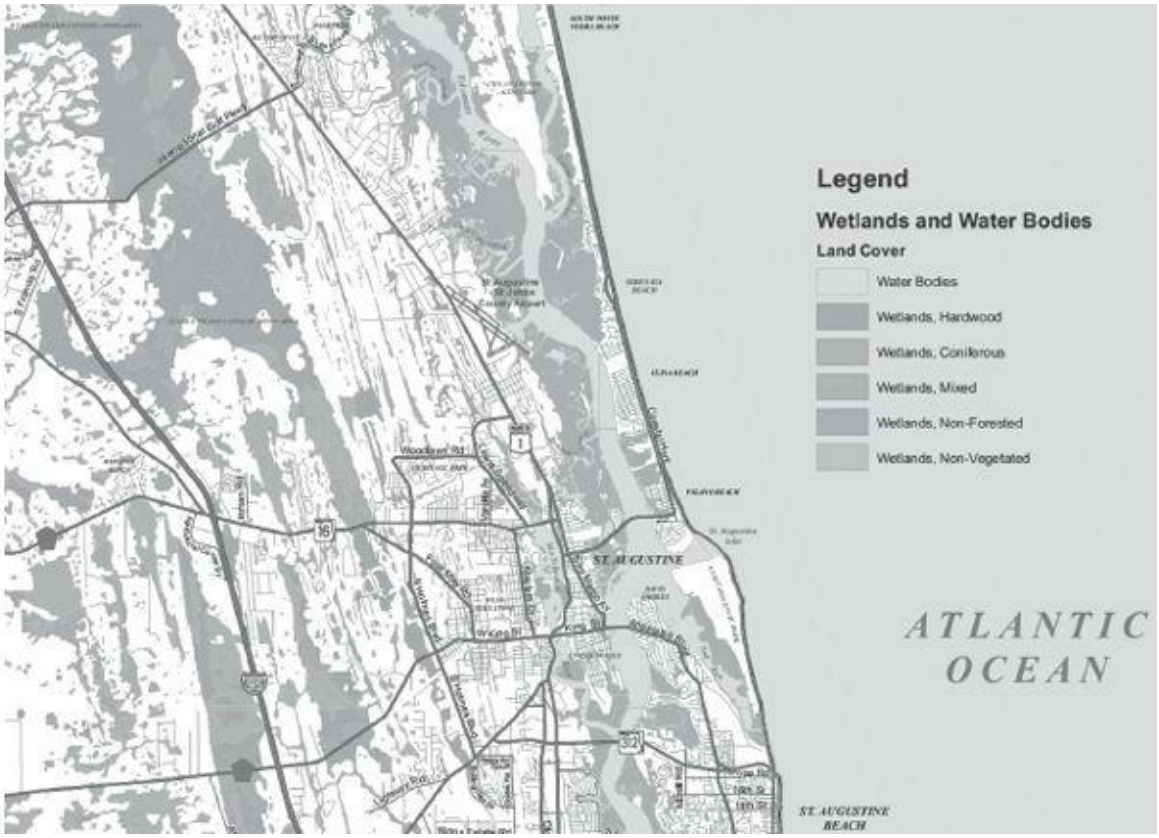


FIGURE 20-11 A portion of the wetlands layer from the St. Johns County, Florida, GIS database.

20-8 Transportation

The Roadway Characteristics Inventory (RCI) is a GIS database which contains information about signs, pavement, drainage structures, bridges, etc., along highways in Florida. First introduced as a

pilot project, the RCI was developed to facilitate pavement maintenance, accident cataloging, replacement or repair of damaged structures (guardrails, signs, fences, etc.), and other items related to highway issues. The GIS replaced manual inventory methods in which facilities were located by driving along the highways and noting the positions of inventory features from odometer readings (dead reckoning). The GIS will be used initially for planning purposes; however, the spatial accuracy is good enough to support other purposes such as preliminary design.

Aerial photography was acquired along the highway at a flying height of 600 m above terrain. This photography was used to compile planimetric features such as road centerlines and edges of pavement. Digital orthophotos were also produced, at a resolution of 10 cm, to provide a spatially accurate visual base map. To complement the photogrammetric data, field inventories were conducted to locate and catalog the various highway features. Field locations were performed using differential, code-phase GPS techniques which obtained positions to submeter accuracy. Figure 20-12 shows a portion of the GIS along a stretch of highway. When a user selects a feature by pointing and clicking, a table pops up, giving detailed information about the particular item. In Fig. 20-12, the box culvert was selected, giving a table showing its position, diameter, type of material, and other information.

FIGURE 20-12 Selection from the Roadway Characteristics Inventory GIS showing relevant information for a specific feature. (*Courtesy 3001, Inc.*)

The plan calls for 5-year updates to be performed to keep the information in the GIS current. Updates will consist of revisiting previously surveyed features to report on their current condition, as well as obtaining information for new features. Due to the high accuracy of the GIS database and its completeness, it is anticipated that the RCI will provide for better decision making now and in the future.

20-9 Multipurpose Land Information System

The city of Sulphur, Louisiana, established a high-accuracy GIS to provide spatial information for a variety of needs. Users of the information include the police, fire, and public works departments. To

achieve high accuracy, the spatial positioning of the GIS was based on GPS control surveys. Additional data layers, created through photogrammetric techniques, were tied to the GPS control. These include a topographic map, digital orthophotography, and a digital elevation model. The system also includes parcel boundaries based on tax assessor data, utility information, stormwater facilities, fire hydrants, and crime incident locations.

This multipurpose GIS supports a variety of uses. For example, if the public works department needs to perform maintenance on a section of sewer pipe, the GIS user can point and click on the pipe and all customers connected to the sewer line will be listed. Notices can then be mailed to these customers to inform them of the impending service interruption. Figure 20-13 shows a portion of the GIS with parcels, buildings, streets, stormwater facilities, water lines, and sanitary sewer lines. A section of sewer pipe was selected, and the table lists the affected addresses. This illustrates just one of the many uses of this multipurpose GIS.

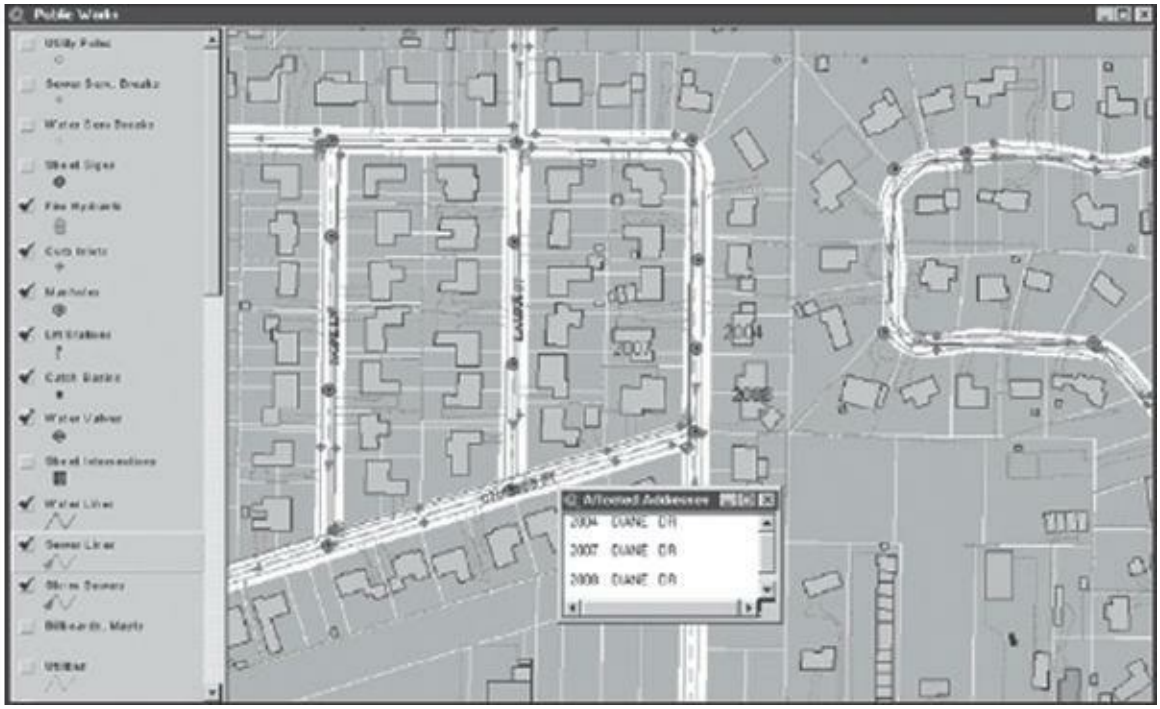


FIGURE 20-13 Sanitary sewer facility map from the Sulphur, Louisiana, GIS showing addresses affected by pipe maintenance. (*Courtesy 3001, Inc.*)

20-10 Summary

The examples presented in this chapter are only a very small sample of the many ways in which GISs are being applied in problem solving. And as indicated by these examples, photogrammetry is being widely used to provide the spatially accurate data layers needed to enable GISs to function. In particular, the examples verify that digital orthophotos are a very commonly used photogrammetric product in GISs. Georeferenced satellite imagery is also frequently used in applications where lower image resolution is suitable. By virtue of their positional accuracies, areawide coverage capabilities, and cost effectiveness, photogrammetric products will continue to be used as data layers in many GIS applications. The references which follow cite numerous other GIS applications where photogrammetry and remote sensing play vital roles.

References

American Society for Photogrammetry and Remote Sensing: "Special Issue: Geographic Information Systems," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 10, 1997.

———: "Special Issue: Remote Sensing and GIS for Hazards," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 10, 1998.

———: "GIS Special Issue: U.S./Mexico Border Region," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 11, 1998.

———: *Manual of Geographic Information Systems*, Bethesda, MD, 2009.

Cadwallader, W., and R. Riethmueller: "Geospatial Technology in a Developing Country," *Geo Info Systems*, vol. 9, no. 1, 1999, p. 22.

Carter, G.: "Estimation of Nonpoint Source Phosphorous and Nitrogen Loads in Five Watersheds in New Jersey's Atlantic Coastal Drainage Basin," *Surveying and Land Information Systems*, vol. 58, no. 3, 1998, p. 167.

Decker, D., and R. Seekins: "Creating a Statewide Digital Base Map: The Texas Orthoimagery Program," *Surveying and Land Information Systems*, vol. 57, no. 1, 1997, p. 23.

Donnelly, J.: "Geographic Information Systems in Map Making," *Bulletin, American Congress on Surveying and Mapping*, no. 134, 1991, p. 30.

Dooley, D.: "Linear Transformation: Delaware Department of Transportation Uses GIS to Fill in the Gaps," *Geo Info Systems*, vol. 8, no. 5, 1998, p. 24.

Douglas, W.: *Environmental GIS: Applications to Industrial Facilities*, Lewis Publishers, Boca Raton, FL, 1995.

Duhaime, R. J., P. V. August, and W. R. Wright: "Automated Vegetation Mapping Using Digital Orthophotography," *Photogrammetric Engineering and Remote Sensing*, vol. 63, no. 11, 1997, p. 1295.

Faber, B. G., W. W. Wallace, and G. E. Johnson: "Active Response GIS: For Resource Management Spatial Decision Support Systems," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 1, 1998, p. 7.

Fry, W., and J. Dozzi: "A New Look at Old Man River," *Civil Engineering*, vol. 67, no. 6, 1997, p. 49.

Gilbrook, M. J.: GIS Paves the Way, *Civil Engineering*, vol. 69, no. 11, 1999, p. 34.

Goldstein, H.: "Mapping Convergence: GIS Joins the Enterprise," *Civil Engineering*, vol. 67, no. 6, 1997, p. 36.

Greenfield, J.: "Consistent Property Line Analysis for Land Surveying and GIS/LIS," *Surveying and Land Information Systems*, vol. 57, no. 2, 1997, p. 69.

Greenwood, D., and D. Hathaway: "Assessing Opal's Impact," *Civil Engineering*, vol. 66, no. 1, 1996, p. 40.

Hendricksen, C., and L. Hall: "GIS Measures Water Use in the Arid West," *Geo Info Systems*, vol. 2, no. 7, 1992, p. 63.

Hinton, C.: "North Carolina City Saves Time, Lives, and Money with Award-Winning GIS," *Geo Info Systems*, vol. 7, no. 9, 1997, p. 35.

Hodgson, M., and R. Palm: "Attitude toward Disaster: A GIS Design for Analyzing Human Response to Earthquake Hazards," *Geo Info Systems*, vol. 2, no. 7, 1992, p. 40.

Kautz, R.: "Satellite Imagery and GIS Help Protect Wildlife Habitat in Florida," *Geo Info Systems*, vol. 2, no. 1, 1992, p. 37.

Koch, M., and F. El-Baz: "Identifying the Effects of the Gulf War on the Geomorphic Features of Kuwait by Remote Sensing and GIS," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 7, 1998, p. 739.

Lang, L.: "Use of GIS, GPS and Remote Sensing Spreads to California's Winegrowers," *Modern Agriculture*, vol. 1, no. 2, 1997, p. 12.
Lembo, A. J., C. Powers, and E. S. Gorin: "The Use of Innovative Data Collection Techniques in Support of Enterprise Wide GIS Development," *Photogrammetric Engineering and Remote Sensing*, vol. 64, no. 9, 1998, p. 861.
Lindquist, R.: "Illinois Cleans Up: Using GIS for Landfill Siting," *Geo Info Systems*, vol. 1, no. 2, 1991, p. 30.
Loukes, D., and J. McLaughlin: "GIS and Transportation: Canadian Perspective," *ASCE Journal of Surveying Engineering*, vol. 117, no. 3, 1991, p. 123.
Lyon, J., and J. McCarthy: *Wetland and Environmental Applications of GIS*, Lewis Publishers, Boca Raton, FL, 1995.
Mettel, C.: "GIS and Satellite Images Provide Precise Calculations of Probable Maximum Flood," *Geo Info Systems*, vol. 2, no. 6, 1992, p. 44.
Oppmann, R., and N. von Meyer: "Oakland County, Michigan, Uses Enterprise GIS to Maintain Its High Quality of Life," *Geo Info Systems*, vol. 9, no. 7, 1999, p. 28.
Padgett, D.: "Assessing the Safety of Transportation Routes for Hazardous Materials," *Geo Info Systems*, vol. 2, no. 2, 1992, p. 46.
Pottle, D.: "Land and Sky," *Point of Beginning*, vol. 23, no. 2, 1997, p. 46.
Quattrochi, D., and J. Luvall: "Urban Sprawl and Urban Pall: Assessing the Impacts of Atlanta's Growth on Meterology and Air Quality Using Remote Sensing and GIS," *Geo Info Systems*, vol. 9, no. 5, 1999, p. 26.
Ripple, W.: *Geographic Information Systems for Resource Management*, American Congress on Surveying and Mapping, Bethesda, MD, 1986.
Voss, D.: "Integrating Pavement Management and GIS," *Geo Info Systems*, vol. 1, no. 9, 1991, p. 18.
Welch, R., M. Madden, and R. F. Doren: "Mapping the Everglades," *Photogrammetric Engineering and Remote Sensing*, vol. 65, no. 2, 1999, p. 163.
Wilkins, D., and J. Beem: "Flight Plan for the Future," *Point of Beginning*, vol. 24, no. 7, 1999, p. 22.

Problems

20-1. Contact the utility service provider in your local area. Describe how it is using GIS to support its operation. Note where photogrammetry is being used to provide data for the GIS.

20-2. Repeat Prob. 20-1, except contact your local tax assessor's office.

20-3. List 10 layers of information which would be useful in a GIS for fire rescue and emergency medical assistance.

20-4. List 10 layers of information which would be useful in a GIS for landfill site selection.

20-5. Give five examples of GIS applications in which digital orthophotos would be useful.

20-6. Cite five examples of GIS applications in which digital elevation models would be useful.

20-7. Give five examples of GIS applications in which topographic data produced by stereocompilation would be useful.

20-8. Browse the National Wetlands Inventory website at www.fws.gov/wetlands and write a brief description of the available products and services.

APPENDIX A

Units, Errors, Significant Figures, and Error Propagation

A-1 Units

The solution of photogrammetric problems generally requires some type of length, angle, or area measurement. Length measurements can be based on either the metric system of meters, centimeters, millimeters, micrometers, etc., or the English system of feet and inches.

While the metric system is preferred for linear units, the English system is still widely used in the United States. Conversion between the two systems is frequently necessary, and it can be complicated by the fact that two common conversion standards exist. Prior to 1959, the accepted metric-English conversion was that 1 m is exactly equal to 39.37 in, or 12 m equals 39.37 ft. This is the basis of the so-called *U.S. survey foot*. In 1959 that conversion was officially changed to 1 in is exactly equal to 2.54 cm or 0.3048 m equals 1 ft, which is the *international standard foot*. The difference between these definitions is approximately 1 part in 500,000 and is therefore negligible in most circumstances. Use of a particular definition is essentially a matter of accepted standards, although in some cases legislative statute dictates the use of a specific conversion. In a practical sense, the only situation of concern occurs in converting geographic coordinates whose values have in excess of six significant figures. In this text, the international standard conversion will be used unless the U.S. survey standard is specifically indicated.

The following list of length, angle, and area unit equivalents should be helpful to students using this book:

1. Length equivalents

a. Metric

1 meter (m) = defined standard
1 micrometer (μm) = 0.000001 m*
1 millimeter (mm) = 0.001 m*
1 centimeter (cm) = 0.01 m*
1 kilometer (km) = 1000 m*

b. English

12 inches (in) = 1 foot (ft)*
3 ft = 1 yard (yd)*
5280 ft = 1 mile (mi)*

c. Metric-English (*international standard*)

2.54 cm = 1 in*
1 m = 3.2808399 ft
0.3048 m = 1 ft*
1 km = 0.6213712 mi

d. Metric-English (U.S. survey foot)

1 m = 39.37 in*
1 m = 3.2808333 ft
0.30480061 m = 1 ft
1 km = 0.6213699 mi

2. Angle equivalents
   $\pi = 3.141592654$
   1 circle = 2π radians (2π rad) = 360 degrees (360°) = 400 grads or gons (400^g)*
   1° = 60 minutes (60′)*
   1′ = 60 seconds (60″)*
   1 rad = 57.29578° = 57°17′44.8″ = 206,264.8″
   1° = 0.01745329 rad = 1.111111^g
   1^g = 0.9° = 54′ = 3240″*
   $1^g = 0.01570796^r$
3. Area equivalents
   a. Metric
      1 hectare (ha) = 10,000 m^2*
      1 km^2 = 100 ha*
   b. English
      1 acre = 43,560 ft^2*
      640 acres = 1 mi^2*
   c. Metric-English (international standard)
      4046.856 m^2 = 1 acre
      1 ha = 2.471054 acres
      0.4046856 ha = 1 acre
   d. Metric-English (U.S. survey foot)
      4046.873 m^2 = 1 acre
      1 ha = 2.471044 acres
      0.4046873 ha = 1 acre

A-2 Errors

In the processes of measuring any quantity, factors such as human limitations, instrumental imperfections, and instabilities in nature render the measured values inexact. Due to these factors, no matter how carefully a measurement is performed, it will always contain some error. Photogrammetry is a science which frequently requires measurements, and therefore an understanding of errors—including how they occur and how they are treated in computation—is important. Before learning about errors, it is helpful to have an understanding of the concepts of *accuracy* and *precision*.

Accuracy can be defined as the degree of conformity to the true value. A value which is very close to the true value has high accuracy, and a value that is far from true has low accuracy. Since the true value for a continuous physical quantity is never known, accuracy is likewise never known; therefore, it can only be estimated. An acceptable method for assessing accuracy is by checking against an independent, higher-accuracy standard.

Precision, on the other hand, is the degree of refinement of a quantity. The level of precision can be assessed by making repeated measurements and checking the consistency of the values. If the values are very close to each other, the measurements have high precision; and if the values vary widely, the measurements have low precision.

An *error* is defined as the difference between a particular value and the true or correct value. Whenever a measurement is made of a continuous physical quantity (such as distance), the result will contain some amount of error. Errors can be categorized as *random errors*, *systematic errors*, and *mistakes*. Mistakes or *blunders* are gross errors caused by carelessness or negligence. Common examples of blunders are point misidentification, a transcription error in recording a value, and

misreading of a scale. Large blunders can generally be avoided or detected and eliminated through careful procedures and subsequent quality control checks. Small blunders can be particularly troublesome when they are large enough to render the results unsatisfactory, yet small enough that they are undistinguishable from acceptable random errors. Blunders can be prevented by exercising care and remaining alert while measurements are being taken. If they occur, they can usually be detected by careful checking and then eliminated.

Systematic error in a measurement is an error which follows some mathematical or physical law. If the conditions causing the error are measured and properly modeled, a correction can be calculated and the systematic error eliminated. Systematic errors will remain constant in magnitude and algebraic sign if the conditions producing them remain the same. Because their algebraic sign tends to remain the same, systematic errors accumulate, and consequently they are often referred to as *cumulative* errors. Examples of systematic errors in photogrammetry are shrinkage or expansion of photographs, camera lens distortions, and atmospheric refraction distortions.

After blunders and systematic errors have been eliminated, the errors that remain are called random or *accidental* errors. Random errors are generally small, but they can never be avoided entirely in measurements. They do not follow physical laws as systematic errors do, and therefore they must be dealt with according to the mathematical laws of probability. Random errors are as likely to be positive as negative; hence they tend to compensate each other and consequently are often called *compensating* errors. Random errors occur in photogrammetry, for example, in estimating between least graduations of a scale, or in indexing a scale.

A-3 Significant Figures

When real numbers are used to represent measured or computed values, it is important to convey the value's precision as well as the number itself. One approach for indicating precision involves enforcement of the rules of *significant figures*. The technically proper method for representing precision involves *statistics*, although significant figures generally yield satisfactory results and are substantially easier to use.

By definition, measured values have a number of significant figures equal to the number of digits that are *certain*, plus one *estimated digit*. As an example, suppose a distance of 24.37 mm is measured with a scale whose smallest graduations are 1/10 mm. The first three digits are certain, and the fourth (7) is estimated on the scale between the 0.3- and 0.4-mm graduations; thus it is also significant. It is important to record *all* significant figures in measurement. Rounding off or failing to record the last estimated digit (even when it is a zero) is a waste of the extra time spent in obtaining that added precision. Conversely, if more than the actual number of significant figures is recorded, a precision is implied which, in fact, does not exist.

While nonzero digits are always significant, zeros in a recorded value may or may not be significant, depending upon the circumstances. The following rules apply:

1. Zeros between nonzero digits are significant.

   Examples:

1001	four significant figures
200.013	six significant figures
4.05	three significant figures

2. Zeros to the left of the first nonzero digit serve only to position the decimal point and are not

significant.

Examples:

0.003 one significant figure
0.057 two significant figures
0.00281 three significant figures

3. Zeros to the right of nonzero digits which are also to the right of a decimal point are significant.

Examples:

0.10 two significant figures
7.50 three significant figures
483.000 six significant figures

4. Zeros to the right of the rightmost nonzero digit but to the left of an implied decimal point are not significant unless specified by placing a bar over the rightmost significant zero or by moving the decimal point to the left and expressing the number in scientific notation. In this book, scientific notation will be used only on rare occasions since it is more difficult to visualize.

Examples:

$380 = 3.8 \times 10^2$ two significant figures
$38\bar{0} = 3.80 \times 10^2$ three significant figures
$160\bar{0} = 1.600 \times 10^3$ four significant figures
$30\bar{0}0 = 3.00 \times 10^3$ three significant figures

In computations using measured values, it is important that answers be given to a number of significant figures consistent with the number of significant figures in the data used to compute them. Specifying an answer to less than the proper number of significant figures does not take advantage of the precision achieved in measuring the quantities, and giving the answer to more significant figures than are justified is misleading because it implies precision that does not exist.

In adding and subtracting, the calculations are performed without regard to significant figures, but the answer is rounded off to the column dictated by the least-precise value. The least-precise value in a set of numbers that are being added or subtracted is that value whose rightmost significant figure is farthest to the left relative to the decimal point.

Examples:

```
  4.735
 +0.67
+24.
 29.405
```

Answer = 29 (rounded to the ones column governed by 24)

```
 1130
 −83.073
 1046.927
```

Answer = 1050 (rounded to the tens column governed by 1130)

In multiplication and division, the number of significant figures in answers is equal to the least

number of significant figures of any data used in the calculation.

Examples:

$1738 \times 24 = 41{,}712$ — Answer = 42,000 (two significant figures governed by 24)

$648.1 \times 0.0523 = 33.89563$ — Answer = 33.9 (three significant figures governed by 0.0523)

$\frac{23.985}{13} = 1.845$ — Answer = 1.8 (two significant figures governed by 13)

In multiplying or dividing by exact constants (or taking reciprocals), the constants do not govern the number of significant figures in the answer.

Examples:

$\frac{1}{209.37} = 0.004776233$ — Answer = 0.0047762 (five significant figures governed by 209.37)

Convert 15.73 ft to inches.

$15.73 \text{ ft} \times 12 \text{ in/ft} = 188.76 \text{ in}$ — Answer = 188.8 in (four significant figures governed by 15.73, not 12)

In mixed calculations, significant figures must be determined by taking into account the order in which the calculations are performed.

Examples:

$$\frac{(104.3 - 101.7)(538)}{104.3} = \frac{2.6(538)}{104.3} = 13.411$$

Answer = 13 (two significant figures governed by 2.6, the result of the subtraction which was performed first)

Solve for i:

$$\frac{1}{152.417} = \frac{1}{30{,}000} = \frac{1}{i}$$

$$0.006560\overline{9}48 = 0.0000\overline{3}333 + \frac{1}{i}$$

$$0.006560\overline{9}48 - 0.0000\overline{3}333 = 0.0065\overline{2}7615 = \frac{1}{i}$$

$i = 153.195$ — Answer = 153 (three significant figures governed by $0.0065\overline{2}7615$, the result of the subtraction)

Note the use of the overbar which serves as a mnemonic to indicate the rightmost significant figure. This avoids rounding of intermediate results, which can contaminate the answer in some cases.

Trigonometric functions involving angles are complicated with regard to assessing the proper number of significant figures. A good method for determining significant figures is to vary the rightmost significant figure in the value by one unit and compute the function a second time, noting the amount of change from the original computed result. This can be readily seen by subtracting the two values and noting the leftmost significant figure in the difference.

Examples:

$\tan(7.50°) = 0.1316525$
$\tan(7.51°) = 0.1318301$
$\text{Difference} = 0.0001776$

Therefore the answer = 0.1317 (round off in the ten-thousandths place, the position of the first significant figure in the difference).

$\cos(0.05°) = 0.999999619$
$\cos(0.06°) = 0.999999452$
$\text{Difference} = 0.000000168$

Therefore the answer = 0.9999996 (round off in the ten-millionth place, the position of the first significant figure in the difference).

$\tan^{-1}(0.457) = 24.5604°$
$\tan^{-1}(0.458) = 24.6078°$
$\text{Difference} = 0.0474°$

Therefore the answer = 24.56° (round off in the hundredths place, the position of the first significant figure in the difference).

In this text the term *nominal* is frequently used to imply nearness to a given value, e.g., a nominal 6-in-focal-length camera. Nominal in this context may be assumed to imply two additional significant figures (a nominal 6-in-focal-length camera lens therefore means a focal length of 6.00 in).

In making computations it is important to carry out intermediate values to more than the required number of significant figures so as not to contaminate the final result. This is easily achieved by storing intermediate values in a calculator's memory storage registers and recalling them as needed. If this is not feasible, carrying one additional digit in intermediate computations should give adequate results.

A-4 Error Propagation

Often it is necessary to estimate the error in a computed quantity which is indirectly determined from measurements that contain error. This can be accomplished through statistical error propagation. Full error propagation assumes that errors in the variables of the equations are *correlated;* i.e., error in one variable is dependent upon error in other variables. Full error propagation is generally not necessary when computing a value as a function of measurements. In its simplest form, error propagation assumes that errors in the measurements are *independent*. Assume that the computed value F, as shown in Eq. (A-1), is a function of n independent variables $x_1, x_2, \ldots, x_n$, which have standard errors $\sigma_1, \sigma_2, \ldots, \sigma_n$, respectively.

$$F = f(x_1, x_2, \ldots, x_n) \tag{A-1}$$

The standard error of the computed value is given by Eq. (A-2).

$$\sigma_F = \pm\sqrt{\left(\frac{\partial F}{\partial x_1}\right)^2 \sigma_1^2 + \left(\frac{\partial F}{\partial x_2}\right)^2 \sigma_2^2 + \cdots + \left(\frac{\partial F}{\partial x_n}\right)^2 \sigma_n^2} \tag{A-2}$$

Example A-1
Compute the value of h and its standard error, based on the following relief displacement equation:

$$h = \frac{dH}{r}$$

where

$$d = 11.3 \text{ mm} \qquad \sigma_d = \pm 0.15 \text{ mm}$$
$$H = 912 \text{ m} \qquad \sigma_H = \pm 3 \text{ m}$$
$$r = 97.2 \text{ mm} \qquad \sigma_r = \pm 0.3 \text{ mm}$$

Solution First, compute h.

$$h = \frac{11.3 \text{ mm} \times 912 \text{ m}}{97.2 \text{ mm}} = 106.02 \text{ m}$$

(Answer = 106 m, based on significant figures)

Next, apply Eq. (A-2).

$$\sigma_h = \pm\sqrt{\left(\frac{\partial h}{\partial d}\right)^2 \sigma_d^2 + \left(\frac{\partial h}{\partial H}\right)^2 \sigma_H^2 + \left(\frac{\partial h}{\partial r}\right)^2 \sigma_r^2}$$

$$= \pm\sqrt{\left(\frac{H}{r}\right)^2 \sigma_d^2 + \left(\frac{d}{r}\right)^2 \sigma_H^2 + \left(\frac{-d \cdot H}{r^2}\right)^2 \sigma_r^2}$$

$$= \pm\sqrt{\left(\frac{912 \text{ m}}{912 \text{ mm}}\right)^2 (0.15 \text{ mm})^2 + \left(\frac{11.3 \text{ mm}}{97.2 \text{ mm}}\right)^2 (3 \text{ m})^2 + \left[\frac{-(11.3 \text{ mm})(912 \text{ m})}{(97.2 \text{ mm})^2}\right]^2 (0.3 \text{ mm})^2}$$

$$= \pm\sqrt{1.98 \text{ m}^2 + 0.12 \text{ m}^2 + 0.11 \text{ m}^2} = \pm 1.5 \text{ m} \quad \blacktriangle$$

References

Anderson, J. M., and E. M. Mikhail: *Surveying: Theory and Practice*, McGraw-Hill, New York, 1998.

Brinker, R. C., and R. Minnick: *The Surveying Handbook*, Chapman & Hall, New York, 1995, chap. 3.

Ghilani, C. D., *Adjustment Computations: Spatial Data Analysis*, 5th ed., Wiley & Sons, Hoboken, NJ, 2010.

Ghilani, C. D., and P. R. Wolf: *Elementary Surveying: An Introduction to Geomatics*, 13th ed., Prentice Hall, NJ, 2012.

Problems

A-1. Convert the following lengths to inches:
(*a*) 23.9 cm
(*b*) 2.00 mm

(*c*) 0.4681 m
(*d*) 22.33 mm

A-2. Convert the following lengths to millimeters:
(*a*) 1.776 in
(*b*) 14.92 in
(*c*) 1.71 ft
(*d*) 0.834 ft

A-3. Make the following length conversions:
(*a*) Express 92.54 m in feet
(*b*) Express 100.00 ft in meters
(*c*) Express 6.67 km in miles
(*d*) Express 6.67 mi in kilometers

A-4. Convert the following angles to degrees, minutes, and seconds.
(*a*) 10.023°
(*b*) 0.050^{g}
(*c*) -8.539^{g}
(*d*) 0.04681^{rad}

A-5. Convert the following angles to grads:
(*a*) 70°03′00″
(*b*) 10°47′37″
(*c*) 11.07°
(*d*) 0.1509^{rad}

A-6. Convert the following angles to radians:
(*a*) 4.245°
(*b*) 15°59′22″
(*c*) 13.435^{g}
(*d*) 0.0074^{g}

A-7. Make the following area conversions.
(*a*) Express 1.9817 m in square feet.
(*b*) Express 9761 ft^2 in hectares.
(*c*) Express 101.0 m^2 in square feet.
(*d*) Express 21,451 ft^2 in square meters.

A-8. How many significant figures are in the following numbers?
(*a*) 2000
(*b*) 43.0
(*c*) 0.006
(*d*) 10,400
(*e*) 3001
(*f*) 42°04′10″

A-9. Express the answers to the following problems to the correct number of significant figures.

(*a*) 76 + 3.17 + 0.5 + 17.125
(*b*) 4.8 × 508 ÷ 95.3
(*c*) The reciprocal of 152.98
(*d*) 161.5 × (135.0 – 133.7) ÷ 212.0
(*e*) $(100.2)^2$
(*f*) $\sqrt{16.001}$
(*g*) The average of 72.32, 72.33, 72.30, and 72.37
(*h*) sin(89.7°)

A-10. Compute the estimated error in h_A based on the following parallax equation:

$$h_A = h_C + \frac{(p_a - p_c)(H - h_C)}{p_a}$$

where

h_C = 54.31 m $\sigma_{h_C} = \pm 0.03$ m
p_a = 98.04 mm $\sigma_{p_a} = \pm 0.05$ mm
p_c = 90.10 mm $\sigma_{p_c} = \pm 0.05$ mm
H = 916.2 m $\sigma_H = \pm 1.0$ m

_____
* The asterisk denotes an exact conversion factor; other factors are correct only to the digits shown.

APPENDIX B

Introduction to Least Squares Adjustment

B-1 Introduction

As discussed in Sec. A-2, all measurements contain error. With appropriate care, blunders can be avoided; and if an appropriate mathematical model is used, compensation for systematic errors can be provided. No matter how much care is taken, however, random errors will still remain. The theoretically correct method of treating these errors is known as *least squares adjustment*. Least squares is by no means a new method. Karl Gauss, a German mathematician, used the method as early as the latter part of the 18th century. Until the invention of computers, however, it was employed rather sparingly because of the lengthy calculations involved. In order to introduce least squares, some fundamental background concepts will be presented first.

B-2 Definitions

The following definitions of terms necessarily must precede a discussion of least squares:

Observations. Directly observed (or measured) quantities which contain random errors.
True value. The theoretically correct or exact value of a quantity. From measurements, however, the true value can never be determined, because no matter how much care is exercised in measurement, small random errors will still always be present.
Error. The difference between any measured quantity and the true value for that quantity. Since the true value of a measured quantity can never be determined, errors are likewise indeterminate, and hence they are strictly theoretical quantities. Errors may be estimated by comparing measured or computed values with those obtained by an independent method known to have higher accuracy. For example, error in a ground distance obtained through scaling from a vertical photograph can be estimated by comparing it with a ground-surveyed value obtained with an electronic distance-measuring device.
Most probable value. That value for a measured or indirectly determined quantity which, based upon the observations, has the highest probability. The most probable value (MPV) is determined through least squares adjustment, which is based on the mathematical laws of probability. The most probable value for a quantity that has been *directly* and *independently* measured several times with observations of equal weight is simply the mean, or

$$\text{MPV} = \frac{\sum_x}{m} \tag{B-1}$$

In Eq. (B-1), Σx is the sum of the individual measurements, and m is the number of observations. Methods for calculating most probable values of quantities determined through *indirect* observations, which may or may not be equally weighted, are described in later sections of this appendix.

Residual. The difference between any measured quantity and the most probable value for that

quantity. It is the value which is dealt with in adjustment computations, since errors are indeterminate. The term *error* is frequently used when *residual* is in fact meant, and although they are very similar, there is this theoretical distinction.

Degrees of freedom. The number of redundant observations (those in excess of the number actually needed to calculate the unknowns). Redundant observations reveal discrepancies in observed values and make possible the practice of least squares adjustment for obtaining most probable values.

Weight. The relative worth of an observation compared to any other observation. Measurements may be weighted in adjustment computations according to their precisions. A very precisely measured value logically should be weighted heavily in an adjustment so that the correction it receives is smaller than that received by a less precise measurement. If the same equipment and procedures are used on a group of measurements, each observation is given an equal weight. Weights are discussed further in Sec. B-6.

Standard deviation. A quantity used to express the precision of a group of measurements. Standard deviation is sometimes called the *root mean square error*, although that designation is somewhat inaccurate. It may also be called the *68 percent error*, since according to the theory of probability, 68 percent of the observations in a group should have residuals smaller than the standard deviation. An expression for the standard deviation of a quantity for which a number of direct, equally weighted observations have been made is

$$S = \pm \frac{\sum v^2}{r} \tag{B-2}$$

In Eq. (B-2), S is the standard deviation, Σv^2 is the sum of the squares of the residuals, and r is the number of degrees of freedom. In the case of *m* repeated measurements of the same unknown quantity (a common occurrence in photogrammetry), the first measurement establishes a value for the unknown, and all additional measurements, $m - 1$ in number, are redundant. The units of standard deviation are the same as those of the original measurements.

Example B-1

The 10 values listed in column (*a*) below were obtained in measuring a photographic distance with a glass scale. Each value was measured using the same instrument and procedures; thus equal weights are assumed. What are the most probable value and the standard deviation of the group of measurements?

(*a*) Measured Values, mm	(*b*) Residuals, mm	(*c*) Squared Residuals, mm²
105.27	−0.005	0.000025
105.26	−0.015	0.000225
105.29	0.015	0.000225
105.29	0.015	0.000225
105.30	0.025	0.000625
105.27	−0.005	0.000025
105.26	−0.015	0.000225
105.28	0.005	0.000025
105.28	0.005	0.000025
105.25	−0.025	0.000625
$\Sigma = 1052.75$	$\Sigma = 0.000$	$\Sigma = 0.002250$

1. By Eq. (B-1), the most probable value for a quantity that has been directly and independently measured several times is

$$\text{MPV} = \frac{1052.75}{10} = 105.275 \text{ mm}$$

2. Residuals listed in column (*b*) above are obtained by subtracting the MPV from each measurement. The squared residuals are listed in column (*c*).
3. By Eq. (B-2), standard deviation is

$$S = \pm\sqrt{\frac{0.002250}{10-1}} = \pm 0.016 \text{ mm} \quad \blacktriangle$$

B-3 Histograms

A *histogram* is a graphical representation of the distribution of a group of measurements or of the residuals for a group of measurements. It illustrates in an "easily digestible" form the nature of occurrence of random errors. A histogram is simply a bar graph of the sizes of measured values, or the sizes of residuals, as abscissas versus their frequency of occurrence as ordinates. An example histogram of residuals for 50 measurements of a photo distance is shown in Fig. B-1.

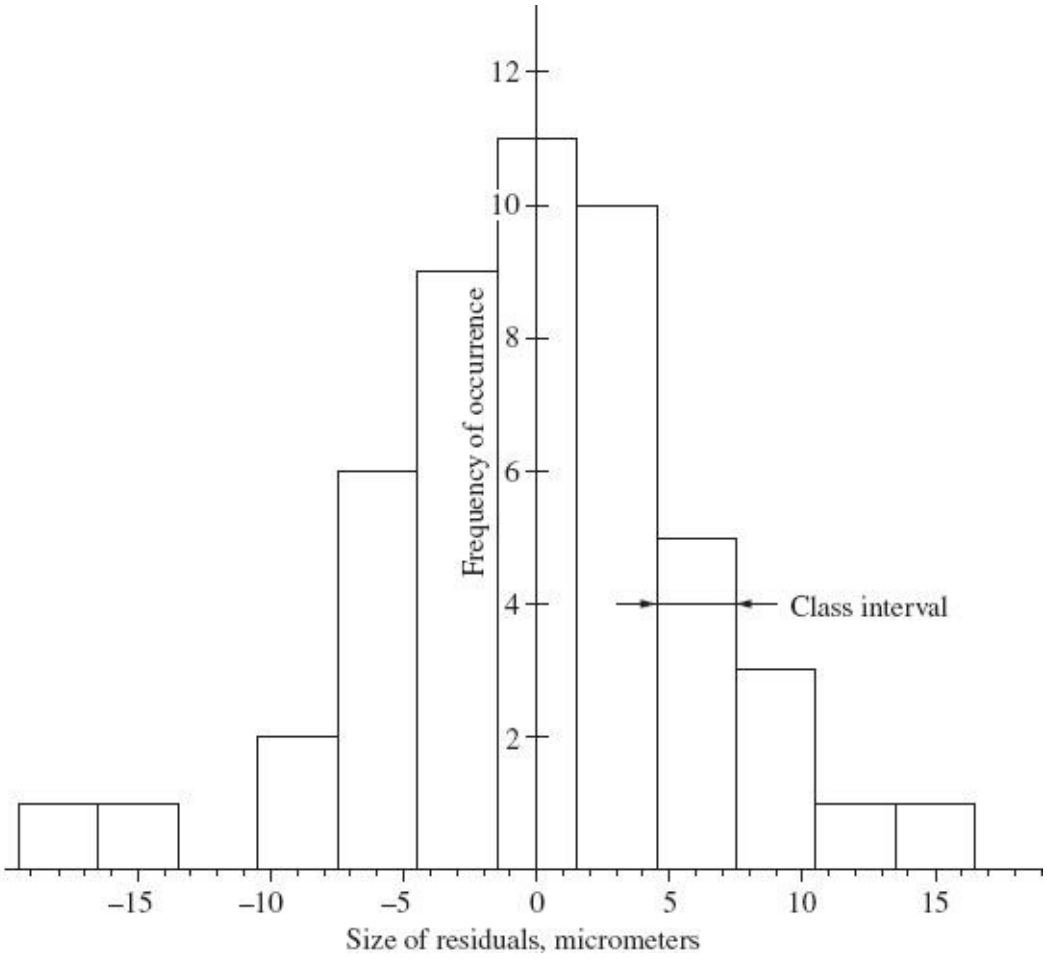


FIGURE B-1 Histogram for 50 measurements of a photographic distance.

Histograms graphically display the following information about a particular group of measurements:

1. Whether the measurements or residuals are symmetrically distributed about some central value.
2. The total spread or dispersion in the measured values or in the residuals.
3. The precision of the measured values. (A tall, narrow histogram represents good precision, while a short, wide one represents poor precision.)

Histograms of varying shape will be obtained by different personnel and for variations in equipment used. Many comparator measurements of a photographic distance, for example, would likely produce a very narrow histogram with a high ordinate at the center. A histogram of the same distance measured the same number of times with an engineer's scale, and plotted at the same scale, would be wider, with a much lower ordinate value at the center.

When a histogram of residuals is plotted, the residuals are calculated and classified into groups or *classes* according to their sizes. The range of residual size in each class is called the *class interval*. The width of bars on a histogram is equal to the class interval, and for the example of Fig. B-1 it is 3 µm. It is important to select a class interval that portrays the distribution of residuals adequately. The value selected will depend upon the total number of measurements in the group. Usually a class interval that produces from 8 to 15 bars (classes) on the graph is ideal. The number of residuals within each class (frequency of residuals) is then counted and plotted on the ordinate scale versus the residual size for the class on the abscissa scale.

B-4 Normal Distribution of Random Errors

In the remaining discussion it is assumed that all error distributions are *normal*. This is a good assumption in photogrammetry since most distributions are in fact normal or very nearly normal.

A normal distribution, or *gaussian distribution* as it is often called, is one in which the graph has the typical bell-shaped curve shown in Fig. B-2. It is symmetric about the ordinate for a residual of zero. For a very large group of measurements, this curve may be obtained in much the same way as a histogram except that the sizes of residuals are plotted on the abscissa versus their *relative frequencies* of occurrence on the ordinate. The relative frequency of residuals occurring within an interval Δv is simply the ratio of the number of residuals within that interval to the total number of residuals. If the number of measurements is infinitely large and if the size of the interval is taken as infinitesimally small, the resulting curve becomes smooth and continuous, as shown in Fig. B-2.

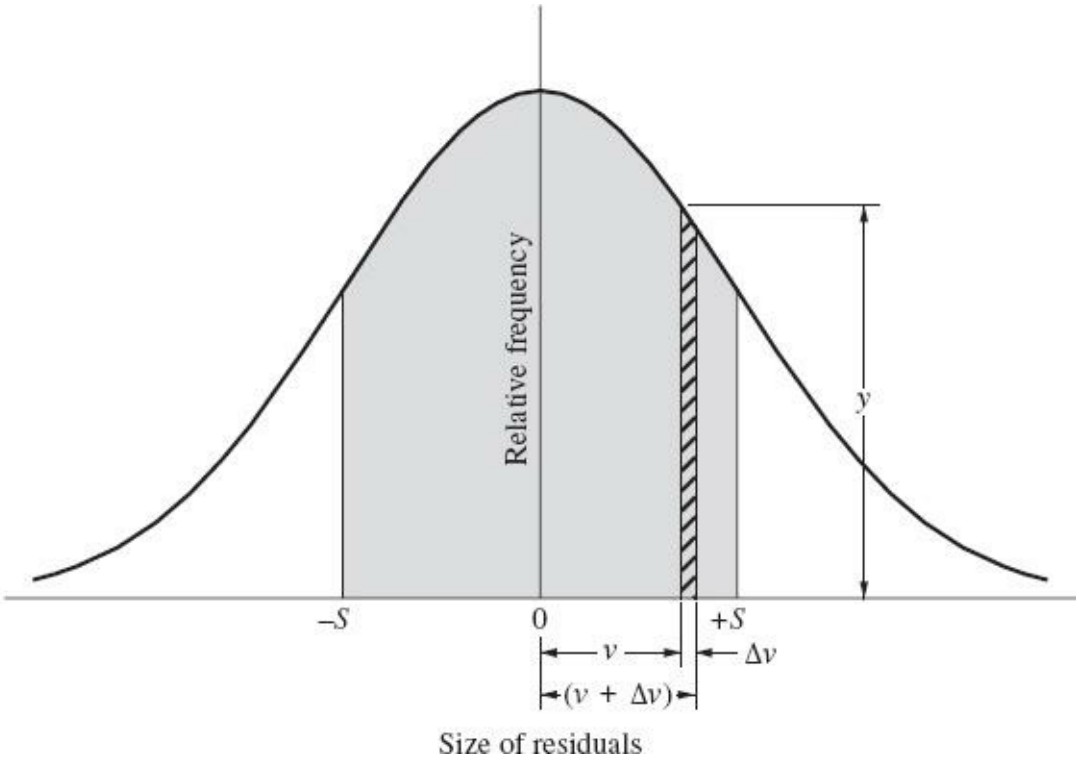


FIGURE B-2 Normal distribution curve.

The equation of the normal distribution curve is

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{B-3}$$

where y = ordinate of normal distribution curve, equal to relative frequency of ccurrence of residuals between the size of v and $v + \Delta v$

σ = standard deviation

e = base of natural logarithms

μ = mean of distribution ($\mu = 0$ in Fig. B-2)

x = abscissa

The derivation of Eq. (B-3) is beyond the scope of this text, but it can be found in references listed at the end of this appendix. In Fig. B-2 the probability of a residual occurring between the limits of v and $v + \Delta v$ is equal to the crosshatched area under the curve between those limits. It is the product of the ordinate y and the interval Δv. For any single measurement of a group of measurements, the probability that its residual occurs between any two abscissas on the curve (such as between –S and +S of Fig. B-2) is equal to the area under the normal distribution curve between those abscissas. Since for a group of measurements all residuals must fall somewhere on the abscissa scale of the normal distribution curve, the total area under the curve represents total probability and is therefore equal to 1.0. The area under the curve between any two abscissas may be found by integrating Eq. (B-3) between those two abscissa limits. The integration is beyond the scope of this text, but it is pertinent to point out that the area between –S and +S (shown as shaded in Fig. B-2) is 68 percent of the total area under the curve. Hence S is called the 68 percent error, as previously mentioned.

B-5 Fundamental Condition of Least Squares

For a group of equally weighted observations, the fundamental condition which is enforced in least squares adjustment is that *the sum of the squares of the residuals is minimized*. This condition, which has been developed from the equation for the normal distribution curve, provides most probable values for the adjusted quantities. Suppose a group of m equally weighted measurements were taken having residuals $v_1, v_2, v_3, \ldots, v_m$. Then, in equation form, the fundamental condition of least squares is expressed as

$$\sum_{i=1}^{m} v_i^2 = v_1^2 + v_2^2 + v_3^2 + \cdots + v_m^2 = \text{minimum} \tag{B-4}$$

Some basic assumptions which underlie least squares theory are that the number of observations being adjusted is large and that the frequency distribution of the errors is *normal*. Although these basic assumptions are not always met, least squares adjustment still provides the most rigorous error treatment available, and hence it has become very popular and important in many areas of modern photogrammetry. Besides yielding most probable values for the unknowns, least squares adjustment enables precisions of adjusted quantities to be determined; and it often reveals the presence of large errors and mistakes so that steps can be taken to eliminate them.

B-6 Weighted Observations

Weights of individual observed values may be assigned according to a priori estimates, or they may be obtained from the standard deviations of the observations, if available. Assuming uncorrelated observations, an equation expressing the relation between standard deviation and weight is

$$w_i = \frac{1}{S_i^2} \tag{B-5}$$

In Eq. (B-5), w_i is the weight of the ith observed quantity, and S_i^2 is the square of the standard deviation, or *variance* of that observation. Equation (B-5) implies that *weight is inversely proportional to variance*. If measured values are to be weighted in least squares adjustment, then the

fundamental condition to be enforced is that the *sum of the weights times their corresponding squared residuals is minimized*, or, in equation form,

$$\sum_{i=1}^{m} w_i v_i^2 = w_1 v_1^2 + w_2 v_2^2 + w_3 v_3^2 + \cdots + w_m v_m^2 = \text{minimum} \tag{B-6}$$

B-7 Applying Least Squares

In the "observation equation" method of least squares adjustment, *observation equations* are written which relate measured values to their residual errors and the unknown parameters. One observation equation is written for each measurement. For a unique solution, the number of equations must equal the number of unknowns. If redundant observations are made, then more observation equations can be written than are needed for a unique solution, and most probable values of the unknown parameters can be determined by the method of least squares. For a group of equally weighted observations, an equation for each residual error is obtained from each observation. The residuals are squared and added to obtain the function $\sum_{i=1}^{m} v_i^2$. To minimize the function, partial derivatives are taken with respect to each unknown variable and set equal to zero. This yields a set of equations called *normal equations* which are equal in number to the number of unknowns. The normal equations are solved to obtain the most probable values for the unknowns.

Example B-2

As an elementary example illustrating the method of least squares adjustment by the observation equation method, consider the following three equally weighted measurements taken between points *A*, *B*, and *C* of Fig. B-3:

$$x + y = 3.0$$
$$x = 1.5$$
$$y = 1.4$$

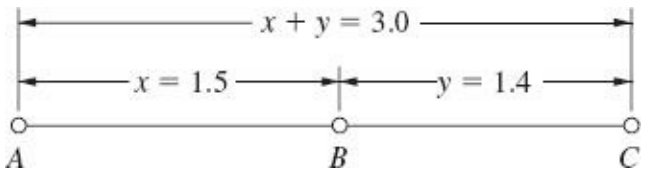


FIGURE B-3 Measurements for least squares adjustment example, Example B-2.

These three equations relate the two unknowns *x* and *y* to the observations. Values for *x* and *y* could be obtained from any two of these equations; therefore, the remaining equation is redundant. Notice, however, that the values obtained for *x* and *y* will differ, depending upon which two equations are solved. It is therefore apparent that the measurements contain errors. The equations may be rewritten as observation equations by including residual errors as follows:

$$x + y = 3.0 + v_1$$
$$x = 1.5 + v_2$$
$$y = 1.4 + v_3$$

To arrive at the least squares solution, the observation equations are rearranged to obtain expressions for the residuals; these are squared and added to form the function $\sum_{i=1}^{m} v_i^2$ as follows:

$$\sum_{i=1}^{m} v_i^2 = (x + y - 3.0)^2 + (x - 1.5)^2 + (y - 1.4)^2$$

The above function is minimized, enforcing the condition of least squares, by taking partial derivatives with respect to each unknown and setting them equal to zero. This yields the following two equations:

$$\frac{\partial \sum v^2}{\partial x} = 0 = 2(x + y - 3.0) + 2(x - 1.5)$$
$$\frac{\partial \sum v^2}{\partial y} = 0 = 2(x + y - 3.0) + 2(y - 1.4)$$

The above equations are the normal equations, and expressed in simplied form they are

$$2x + y = 4.5$$
$$x + 2y = 4.4$$

Solving the simplied normal equations simultaneously yields x = 1.533 and y = 1.433. According to the theory of least squares, these values have the highest probability. Having the most probable values for the unknowns, the residuals can be calculated by substitution back into the original observation equations, or

$$v_1 = 1.533 + 1.433 - 3.000 = -0.033$$
$$v_2 = 1.533 - 1.500 = 0.033$$
$$v_3 = 1.433 - 1.400 = 0.033 \quad \blacktriangle$$

The above example is indeed simple, but it serves to illustrate the method of least squares without complicating the mathematics. Least squares adjustment of large systems of observations is performed in the same manner.

B-8 Systematic Formulation of Normal Equations

In large systems of observations, it is helpful to utilize systematic procedures to formulate normal equations. Consider the following system of *m* linear observation equations of equal weight containing *n* unknowns:

$$
\begin{aligned}
a_{11}X_1 + a_{12}X_2 + a_{13}X_3 + \cdots + a_{1n}X_n - L_1 &= v_1 \\
a_{21}X_1 + a_{22}X_2 + a_{23}X_3 + \cdots + a_{2n}X_n - L_2 &= v_2 \\
a_{31}X_1 + a_{32}X_2 + a_{33}X_3 + \cdots + a_{3n}X_n - L_3 &= v_3 \\
\cdots\cdots\cdots

The formulation of normal equations from observation equations may be further systematized by handling the systems of Eqs. (B-8) or (B-9) in a tabular manner.

Example B-3
Using the tabular method, form the normal equations for Example B-2.

Solution

Equation Number (i)	a_{i1}	a_{i2}	L_i	$a_{i1}a_{i1}$	$a_{i1}a_{i2}$	$a_{i2}a_{i2}$	$a_{i1}L_i$	$a_{i2}L_i$
1	1	1	3.0	1	1	1	3.0	3.0
2	1	0	1.5	1	0	0	1.5	0.0
3	0	1	1.4	0	0	1	0.0	1.4
				$\Sigma = 2$	$\Sigma = 1$	$\Sigma = 2$	$\Sigma = 4.5$	$\Sigma = 4.4$

For this example, normal equations are formed by satisfying Eqs. (B-8) as follows (note that because of the commutative property of multiplication, $a_{i1}a_{i2} = a_{i2}a_{i1}$.):

$$\sum_{i=1}^{m}(a_{i1}a_{i1})X_1 + \sum_{i=1}^{m}(a_{i1}a_{i2})X_2 = \sum_{i=1}^{m}(a_{i1}L_i)$$

$$\sum_{i=1}^{m}(a_{i2}a_{i1})X_1 + \sum_{i=1}^{m}(a_{i2}a_{i2})X_2 = \sum_{i=1}^{m}(a_{i2}L_i)$$

Substituting the appropriate values from the above table yields the required normal equations as follows:

$$2x + y = 4.5$$
$$x + 2y = 4.4 \quad \blacktriangle$$

B-9 Matrix Methods in Least Squares Adjustment

It has been previously mentioned that least squares computations are quite lengthy and are therefore most economically performed on a computer. The algebraic approach—Eqs. (B-8) and (B-9)—for forming normal equations, and for obtaining their simultaneous solution, can be programmed for computer solution or set up in a computer spreadsheet. The procedure is much more easily adapted to matrix methods, however. Readers may wish to consult references listed at the end of this appendix for a review of matrix terminology and operations.

In developing matrix equations for least squares computations, analogy will be made to the algebraic approach presented in Sec. B-8. First, observation Eqs. (B-7) may be represented in matrix form as

$$_mA^n\,{}_nX^1 = {}_mL^1 + {}_mV^1 \qquad \text{(B-10)}$$

where

$$ {}_mA^n = \begin{bmatrix} a_{11} & a_{21} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ a_{41} & a_{42} & a_{43} & \cdots & a_{4n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \qquad {}_nX^1 = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_n \end{bmatrix} $$

$$ {}_mL^1 = \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ \vdots \\ L_m \end{bmatrix} \qquad {}_mV^1 = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ \vdots \\ V_m \end{bmatrix} $$

Upon studying the following matrix representation, it will be noticed that normal Eqs. (B-8) are obtained as follows:

$$ (A^TA)X = A^TL \tag{B-11} $$

In the above equation, A^TA is the matrix of normal equation coefficients of the unknowns. Premultiplying both sides of Eq. (B-11) by $(A^TA)^{-1}$ and reducing, results in:

$$ \begin{aligned} (A^TA)^{-1}\,(A^TA)X &= (A^TA)^{-1}(A^TL) \\ IX &= (A^TA)^{-1}(A^TL) \\ X &= (A^TA)^{-1}(A^TL) \end{aligned} \tag{B-12} $$

In the above reduction, I is the identity matrix. Equation (B-12) is the basic least squares matrix equation for *equally weighted* observations. The matrix X consists of most probable values for unknowns X_1, X_2, X_3, …, X_n. For a system of *weighted* observations, the following matrix equation provides the X matrix of most probable values for the unknowns:

$$ X = (A^TWA)^{-1}(A^TWL) \tag{B-13} $$

In Eq. (B-13) the matrices are identical to those of the equally weighted equations, with the inclusion of the W matrix, which is a diagonal matrix of weights and is defined as follows:

$$W = \begin{bmatrix} w_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & w_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & w_3 & 0 & \cdots & 0 \\ 0 & 0 & 0 & w_4 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & w_m \end{bmatrix}$$

In the preceding W matrix, all off-diagonal elements are shown as zeros. This is proper when the individual observations are independent and uncorrelated; i.e., they are not dependent upon each other. This is often the case in photogrammetric applications.

Example B-4

Solve Example B-2 by using matrix methods.

The observation equations of Example B-2 may be expressed in matrix form as follows:

$$_3A^2 \, _2X^1 = _3L^1 + _3V^1$$

where

$$_3A^2 = \begin{bmatrix} 11 \\ 10 \\ 01 \end{bmatrix} \quad _2X^1 = \begin{bmatrix} x \\ y \end{bmatrix} \quad _3L^1 = \begin{bmatrix} 3.0 \\ 1.5 \\ 1.4 \end{bmatrix} \quad _3V^1 = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

Solving matrix Eq. (B-12) gives

$$A^T A = \begin{bmatrix} 110 \\ 101 \end{bmatrix} \begin{bmatrix} 11 \\ 10 \\ 01 \end{bmatrix} = \begin{bmatrix} 21 \\ 12 \end{bmatrix}$$

$$(A^T A)^{-1} = \begin{bmatrix} 0.6667 & -0.3333 \\ -0.3333 & 0.6667 \end{bmatrix} A^T L = \begin{bmatrix} 4.5 \\ 4.4 \end{bmatrix}$$

$$X = (A^T A)^{-1}(A^T A) = \begin{bmatrix} 0.6667 & -0.3333 \\ -0.3333 & 0.6667 \end{bmatrix} \begin{bmatrix} 4.5 \\ 4.4 \end{bmatrix} = \begin{bmatrix} 1.533 \\ 1.433 \end{bmatrix}$$

Note that this solution yields exactly the same values for x and y as were obtained through the algebraic approach of Example B-2. ▲

B-10 Matrix Equations for Precisions of Adjusted Quantities

The matrix equation for calculating residuals after adjustment, whether the adjustment is weighted or not, is

$$V = AX - L \tag{B-14}$$

The standard deviation of unit weight for an unweighted adjustment is

$$S_0 = \sqrt{\frac{V^T V}{r}} \tag{B-15}$$

The standard deviation of unit weight for a weighted adjustment is

$$S_0 = \sqrt{\frac{V^T W V}{r}} \tag{B-16}$$

I n Eqs. (B-15) and (B-16), r is the number of degrees of freedom and equals the number of observation equations minus the number of unknowns, or $r = m - n$.

Standard deviations of the adjusted quantities are

$$S_{x_i} = S_0 \sqrt{Q_{X_i X_i}} \tag{B-17}$$

In Eq. (B-17), S_{x_i} is the standard deviation of the ith adjusted quantity, i.e., the quantity in the ith row of the X matrix; S_0 is the standard deviation of unit weight as calculated by Eq. (B-15) or (B-16); and $Q_{x_i x_i}$. is the element in the ith row and the ith column of the matrix $(A^T A)^{-1}$ in the unweighted case, or the matrix $(A^T W A)^{-1}$ in the weighted case. The matrices $S_0^2 (A^T A)^{-1}$ and $S_0^2 (A^T W A)^{-1}$ are the *covariance* matrices.

Example B-5

Calculate the standard deviation of unit weight and the standard deviations of the adjusted quantities x and y for the unweighted problem of Example B-4.

(a) By Eq. (B-14), the residuals are

$$V = \begin{bmatrix} 11 \\ 10 \\ 01 \end{bmatrix} \begin{bmatrix} 1.533 \\ 1.433 \end{bmatrix} - \begin{bmatrix} 3.0 \\ 1.5 \\ 1.4 \end{bmatrix} = \begin{bmatrix} -0.033 \\ 0.033 \\ 0.033 \end{bmatrix}$$

(b) By Eq. (B-15), the standard deviation of unit weight is

$$V^T V = \begin{bmatrix} -0.033 & 0.033 & 0.033 \end{bmatrix} \begin{bmatrix} -0.033 \\ 0.033 \\ 0.033 \end{bmatrix} = 0.0033$$

$$S_0 = \sqrt{\frac{0.0033}{3-2}} = \pm 0.057$$

(c) Using Eq. (B-17), the standard deviations of the adjusted values for x and y are calculated as

$$S_x = \pm 0.057\sqrt{0.6667} = \pm 0.047$$

$$S_y = \pm 0.057\sqrt{0.6667} = \pm 0.047$$

In part (*c*) above, the numbers 0.6667 within the radicals are the (1,1) and (2,2) elements of the $(A^TA)^{-1}$ matrix of Example B-4. The interpretation of the standard deviations computed under part (*c*) is that there is a 68 percent probability that the true values for *x* and *y* are within ±0.047 of their adjusted values. Note that for this simple example the magnitudes of the three residuals calculated in part (*a*) were equal, and that the standard deviations of *x* and *y* were equal in part (*c*). This is due to the symmetric nature of this particular problem (illustrated in Fig. B-3), but this is seldom, if ever, the case with more complex problems. ▲

B-11 Practical Example

The following example is presented to illustrate a practical application of least squares in photogrammetry. The example also shows the method of calculating the coefficients of a polynomial which approximates the symmetric radial lens distortion curve for an aerial camera (see Sec. 4-11).

Example B-6

From the calibration data of an aerial camera given in the table below, calculate the coefficients of a polynomial which models the symmetric radial lens distortion curve.

Radial Distance *r*, mm	Radial Lens Distortion Δ*r*, mm
20.170	0.004
41.051	0.007
63.460	0.007
88.454	0.001
107.276	−0.003
128.555	−0.004

Solution As presented in Sec. 4-11, a polynomial of the following form is the appropriate model for symmetric radial lens distortion:

$$\Delta r = k_1 r + k_2 r^3 + k_3 r^5 + k_4 r^7 \tag{B-18}$$

In Eq. (B-18), Δ*r* is the symmetric radial lens distortion at a radial distance *r* from the principal point. The *k*'s are coefficients which define the shape of the distortion curve. One equation of the form of Eq. (B-18) can be written for each radial distance at which the distortion is known from calibration. Since there are four *k*'s, four equations are required to obtain a unique solution for them. From calibration, distortions are determined for six radial distances; hence six equations can be written, and the *k*'s may be computed by least squares. In polynomial equations of this type, there is commonly a problem with ill-conditioning of the normal equations due to having numbers raised to large powers. To reduce this problem, it is useful to convert the radial distances to meters so that when raised to the seventh power, the result is less than 1. This is not a foolproof approach, although it is effective in

many cases.

Based on the calibration data, the following observation equations, in the matrix form of Eq. (B-10), may be written (note that the radial distances have been converted to meters):

$$\begin{bmatrix} (0.020170) & (0.020170)^3 & (0.020170)^5 & (0.020170)^7 \\ (0.041051) & (0.041051)^3 & (0.041051)^5 & (0.041051)^7 \\ (0.063460) & (0.063460)^3 & (0.063460)^5 & (0.063460)^7 \\ (0.088454) & (0.088454)^3 & (0.088454)^5 & (0.088454)^7 \\ (0.107276) & (0.107276)^3 & (0.107276)^5 & (0.107276)^7 \\ (0.128555) & (0.128555)^3 & (0.128555)^5 & (0.128555)^7 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} = \begin{bmatrix} 0.004 \\ 0.007 \\ 0.007 \\ 0.001 \\ -0.003 \\ -0.004 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix}$$

The least squares solution in the form of Eq. (B-11) gives

$$\begin{bmatrix} 4.19778\times10^{-2} & 4.85999\times10^{-4} & 6.58695\times10^{-6} & 9.61536\times10^{-8} \\ 4.85999\times10^{-4} & 6.58695\times10^{-6} & 9.61536\times10^{-8} & 1.46503\times10^{-9} \\ 6.58695\times10^{-6} & 9.61536\times10^{-8} & 1.46503\times10^{-9} & 2.29302\times10^{-11} \\ 9.61536\times10^{-8} & 1.46503\times10^{-9} & 2.29302\times10^{-11} & 3.65246\times10^{-13} \end{bmatrix}$$

$$X = \begin{bmatrix} 6.46630\times10^{-5} \\ -9.20374\times10^{-6} \\ -1.69618\times10^{-7} \\ -2.73878\times10^{-9} \end{bmatrix}$$

Solving this equation for the unknowns gives the following:

$$X = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} = \begin{bmatrix} 0.229581 \\ -35.8926 \\ 1{,}018.26 \\ 12{,}105.0 \end{bmatrix}$$

By using these k's in Eq. (B-18), estimates for radial lens distortions (in millimeters) for any value of r (in meters) may be readily calculated. ▲

References

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Benjamin, J. R., and C. A. Cornell: *Probability, Statistics and Decision for Civil Engineers*, McGraw-Hill, New York, 1970.

Bhattacharyya, G. K., and R. A. Johnson: *Statistical Concepts and Methods*, Wiley & Sons, New York, 1977.

Crandall, K. C., and R. W. Seabloom: *Engineering Fundamentals in Measurements, Probability and Dimensions*, McGraw-Hill, New York, 1970.

Ghilani, C. D., *Adjustment Computations: Spatial Data Analysis*, 5th ed., Wiley & Sons, Hoboken, NJ, 2010.

Hardy, R. L.: "Least Squares Prediction," *Photogrammetric Engineering and Remote Sensing*, vol. 43, no. 4, 1977, p. 475.

Hirvonen, R. A.: *Adjustment by Least Squares in Photogrammetry and Geodesy*, Frederick Ungar Publishing, New York, 1971.
Mikhail, E. M.: "Parameter Constraints in Least Squares," *Photogrammetric Engineering*, vol. 36, no. 12, 1970, p. 1277.
———: *Observations and Least Squares*, Harper & Row, New York, 1976.
Rampal, K. K.: "Least Squares Collocation in Photogrammetry," *Photogrammetric Engineering and Remote Sensing*, vol. 42, no. 5, 1976, p. 659.
Wong, K. W.: "Propagation of Variance and Covariance," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 1, 1975, p. 75.
Zimmerman, D. S.: "Least Squares by Diagonal Partitioning," *Canadian Surveyor*, vol. 28, no. 5, 1974, p. 677.

Problems

B-1. A photogrammetric distance was measured 10 times using the same equipment and procedures with the following results: 76.74, 76.69, 76.70, 76.72, 76.69, 76.75, 76.72, 76.77, 76.70, and 76.71 mm. Calculate the most probable value for the photo distance and the standard deviation of the group of measurements.

B-2. Repeat Prob. B-1, except that the following 15 measurements were obtained: 44.29, 44.37, 44.40, 44.38, 44.43, 44.37, 44.39, 44.39, 44.42, 44.38, 44.44, 44.40, 44.37, 44.41, and 44.42 mm.

B-3. Compute the most probable values of unknowns x_1, x_2, and x_3 for the following observation equations, using a computer spreadsheet, and calculate the standard deviations of the adjusted quantities.

$$2x_1 + 3x_2 + x_3 = 15$$
$$x_1 - 2x_2 + 3x_3 = -12$$
$$7x_1 + x_2 - 2x_3 = 23$$
$$-x_1 - x_2 - x_3 = -4$$

B-4. Suppose the constant terms 15, –12, 23, and –4 of the four equations of Prob. B-3 represent measurements having relative weights of 2, 3, 1, and 4, respectively. Using weighted least squares, calculate most probable values for x_1, x_2, and x_3 and determine the standard deviations of these values.

B-5. Repeat Prob. B-3, except that the four equations are as follows:

$$5x_1 + 3x_2 - 2x_3 = -20$$
$$2x_1 - x_2 + 6x_3 = 23$$
$$x_1 + 3x_2 + x_3 = 12$$
$$-4x_1 + 3x_2 - 3x_3 = 5$$

B-6. If the constant terms –20, 23, 12, and 5 of Prob. B-5 represent measurements having relative weights of 4, 2, 1, and 2, respectively, calculate the least squares solution for the unknowns and

determine their standard deviations.

APPENDIX C

Coordinate Transformations

C-1 Introduction

A problem frequently encountered in photogrammetric work is conversion from one rectangular coordinate system to another. This is because photogrammetrists commonly determine coordinates of unknown points in convenient *arbitrary* rectangular coordinate systems. These arbitrary coordinates may be read from comparators or stereoscopic plotters, or they may result from analytic computation. The arbitrary coordinates must then be converted to a *final* system, such as the camera photo coordinate system in the case of comparator measurements, or to a ground coordinate system, such as the state plane coordinate system in the case of stereoplotter or analytically derived arbitrary model coordinates. The procedure for converting from one coordinate system to another is known as *coordinate transformation*. The procedure requires that some points have their coordinates known (or measured) in both the arbitrary and the final coordinate systems. Such points are called *control points*. Figure C-1 shows the effects of applying three common transformation types.

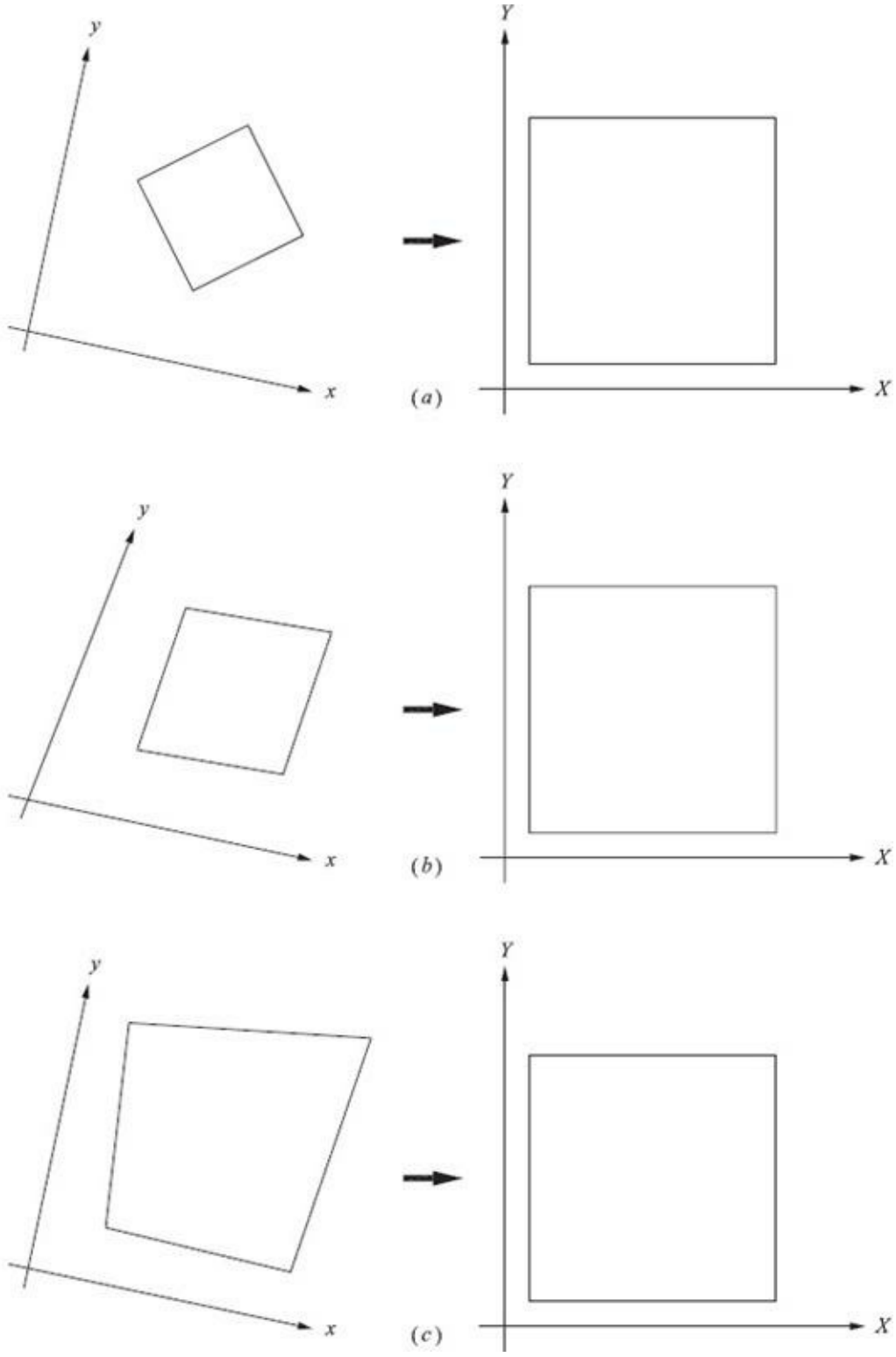


FIGURE C-1 Effects of different coordinate transformations. (*a*) Conformal coordinate transformation. (*b*) Affine coordinate transformation. (*c*) Projective coordinate transformation.

C-2 Two-Dimensional Conformal Coordinate Transformation

The term *two-dimensional* means that the coordinate systems lie on plane surfaces. A *conformal* transformation is one in which true shape is preserved after transformation. To perform a two-dimensional conformal coordinate transformation, it is necessary that coordinates of at least two points be known in both the arbitrary and final coordinate systems. Accuracy in the transformation is improved by choosing the two points as far apart as possible. If more than two control points are

available, an improved solution may be obtained by applying the method of least squares.

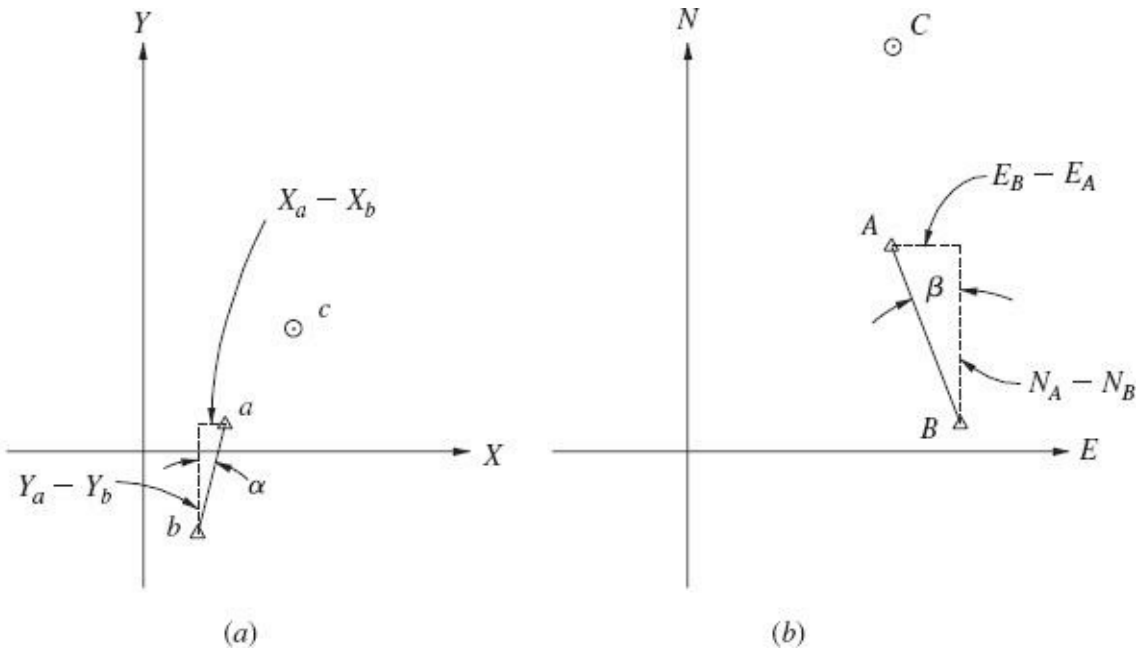


FIGURE C-2 (*a*) Arbitrary *XY* two-dimensional coordinate system. (*b*) Ground *EN* two-dimensional system.

A two-dimensional conformal coordinate transformation consists of three basic steps: (1) scale change, (2) rotation, and (3) translation. The example illustrated in Fig. C-2 is used to demonstrate the procedure. This example uses the minimum of two control points. Section C-4 describes the procedure when more than two control points are available. Figure C-2*a* shows the positions of points *a* through *c*, whose coordinates are known in an arbitrary *XY* system. Figure C-2*b* illustrates the positions of the same points, labeled *A* through *C* in a (ground) *EN* system. The coordinates of *A* and *B* are known in the ground system, and it is required to determine the coordinates of *C* in the ground system.

Step 1: Scale Change

By comparing Figs. C-2*a* and *b*, it is evident that the lengths of lines *ab* and *AB* are unequal, hence the scales of the two coordinate systems are unequal. The scale of the *XY* system is made equal to that of the *EN* system by multiplying each *X* and *Y* coordinate by a scale factor *s*. The scaled coordinates are designated as *X′* and *Y′*. By use of the two control points, the scale factor is calculated in relation to the two lengths *AB* and *ab* as

$$s = \frac{AB}{ab} = \frac{\sqrt{(E_B - E_A)^2 + (N_B - N_A)^2}}{\sqrt{(X_b - X_a)^2 + (Y_b - Y_a)^2}} \tag{C-1}$$

Step 2: Rotation

If the scaled *X′Y′* coordinate system is superimposed over the *EN* system of Fig. C-2*b* so that line *AB* in both systems coincides, the result is as shown in Fig. C-3. An auxiliary axis system *E′N′* is constructed through the origin of the *X′Y′* axis system parallel to the *EN* axes. It is necessary to rotate from the *X′Y′* system to the *E′N′* system, or in other words, to calculate *E′N′* coordinates for the unknown points from their *X′Y′* coordinates. The *E′N′* coordinates of point *C* may be calculated in

terms of the clockwise angle θ by using the following equations:

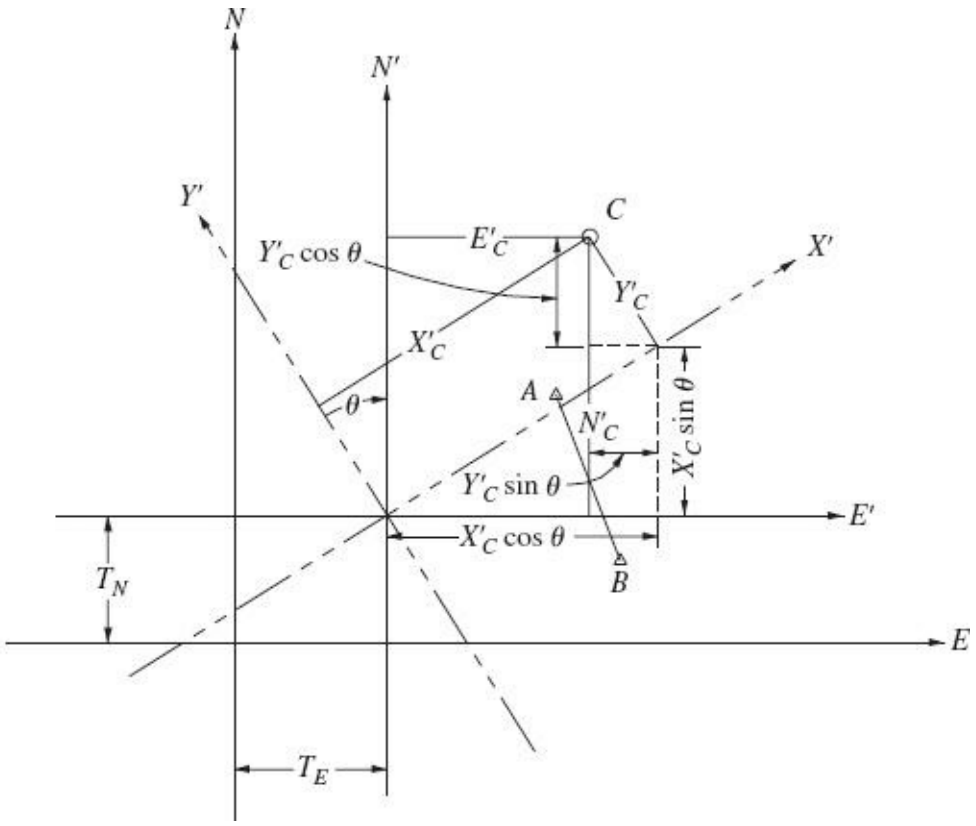


FIGURE C-3 Scaled $X'Y'$ coordinate system superimposed onto the EN ground coordinate system.

$$E'_C = X'_C \cos\theta - Y'_C \sin\theta$$
$$N'_C = X'_C \sin\theta - Y'_C \cos\theta \tag{C-2}$$

Rotation angle θ, shown in Fig. C-3, is the sum of angles α and β which are indicated on Figs. C-2*a* and *b*. From coordinates of the two control points, these angles are calculated as

$$\alpha = \tan^{-1}\left(\frac{X_a - X_b}{Y_a - Y_b}\right)$$
$$\beta = \tan^{-1}\left(\frac{E_B - E_A}{N_A - N_B}\right) \tag{C-3}$$

Step 3: Translation

The final step in the coordinate transformation is a translation of the origin of the $E'N'$ system to the origin of the EN system. The translation factors required are T_E and T_N, which are illustrated in Fig. C-3. Final E and N ground coordinates for points C then are

$$E_C = E'_C + T_E$$
$$N_C = N'_C + T_N \tag{C-4}$$

where the translation factors T_E and T_N are calculated as

$$T_E = E_A - E'_A = E_B - E'_B$$
$$T_N = N_A - N'_A = N_B - N'_B \tag{C-5}$$

Note from Eqs. (C-5) that these translation factors may be calculated in two different ways by using either control point *A* or *B*. It is advisable to calculate them by using both points, to obtain a computation check.

Working sketches are recommended in computing coordinate transformations to aid in reducing the likelihood of mistakes. Caution should be exercised to ensure that correct algebraic signs are applied to the coordinates used in the transformation equations.

Example C-1

Assume that in Figs. C-2*a* and *b* the arbitrary and ground coordinates of points *A* through *C* are as follows:

	Arbitrary Coordinates		Ground Coordinates	
Point	***X***	***Y***	***E***	***N***
A	632.17	121.45	1100.64	1431.09
B	355.20	−642.07	1678.39	254.15
C	1304.81	596.37		

It is required to compute the coordinates of point *C* in the ground *EN* system.

Solution

(*a*) The scale factor is calculated from Eq. (C-1) as

$$s = \frac{\sqrt{(1678.39 - 1100.64)^2 + (254.15 - 1431.09)^2}}{\sqrt{(355.20 - 632.17)^2 + (-642.07 - 121.45)^2}}$$
$$= \frac{1311.10}{812.20} = 1.61425$$

The arbitrary coordinates are then expanded to the *X′Y′* system, which is equal in scale to the ground coordinate system, by multiplying each of the arbitrary coordinates by the scale factor. After multiplication, the *X′Y′* coordinates are as follows:

	Scaled Coordinates	
Point	***X′***	***Y′***
A	1020.48	196.05
B	573.38	−1036.46
C	2106.29	962.69

(*b*) The rotation angle, calculated by Eqs. (C-3), is

$$\tan\alpha = \frac{632.17 - 355.20}{121.45 + 642.07} = 0.362754 \qquad \alpha = 19.9384^\circ$$

$$\tan\beta = \frac{1678.39 - 1100.64}{1431.09 + 254.15} = 0.490892 \qquad \beta = 26.1460^\circ$$

$$\theta = 19.9384^\circ + 26.1460^\circ = 46.0845^\circ$$

Rotation Eqs. (C-2) are then solved to obtain E' and N' coordinates. The solution in tabular form is as follows (with sin θ = 0.720363, and cos θ = 0.693597):

Point	$X'\cos\theta$	$Y'\sin\theta$	$X'\sin\theta$	$Y'\cos\theta$	E'	N'
A	707.80	141.23	735.12	135.98	566.57	871.10
B	397.70	−746.63	413.04	−718.89	1144.32	−305.84
C	1460.92	693.49	1517.29	667.72	767.43	2185.01

(*c*) The translation factors T_E and T_N are calculated next, using Eqs. (C-5) as follows:

$$T_E = E_A - E_A' = 1100.64 - 566.58 = 534.07$$

also $$T_E = E_B - E_B' = 1678.39 - 1144.32 = 534.07 \qquad \textit{Check!}$$

$$T_N = N_A - N_A' = 1431.09 - 871.10 = 559.99$$

and $$T_N = N_B - N_B' = 254.15 + 305.84 = 559.99 \qquad \textit{Check!}$$

By adding the translation factors to E_C' and N_C', the following final transformed E and N coordinates are obtained for point *C*:

$$E_C = 767.43 + 534.07 = 1301.49$$

$$N_C = 2185.01 + 559.99 = 2745.01 \quad \blacktriangle$$

In the above example, although only one unknown point (point *C*) was transformed, any number of points could have been transformed by using just the two control points.

C-3 Alternative Method of Two-Dimensional Conformal Coordinate Transformation

When a computer is used, it is advantageous to compute coordinate transformations by an alternative method. In this method, equations involving four transformation coefficients are formulated in terms of the coordinates of the two or more points whose positions are known in both coordinate systems. The formation of the equations follows the same three steps discussed in Sec. C-2. The procedure consists of first multiplying each of the original coordinates of points *a* and *b* by a scale factor. The following four equations result:

$$X'_a = sX_a$$
$$Y'_a = sY_a$$
$$X'_b = sX_b$$
$$Y'_b = sY_b$$

(C-6)

Equations (C-6) are now substituted into Eqs. (C-2), except that the subscripts of Eqs. (C-2) are changed to be applicable for points *A* and *B*. This substitution yields

$$E'_A = sX_a \cos\theta - sY_a \sin\theta$$
$$N'_A = sX_a \sin\theta + sY_a \cos\theta$$
$$E'_B = sX_b \cos\theta - sY_b \sin\theta$$
$$N'_B = sX_b \sin\theta + sY_b \cos\theta$$

(C-7)

Finally, translation factors T_E and T_N, as described previously, are added to Eqs. (C-7) to yield the following equations:

$$E_A = sX_a \cos\theta - sY_a \sin\theta + T_E$$
$$N_A = sX_a \sin\theta + sY_a \cos\theta + T_N$$
$$E_B = sX_b \cos\theta - sY_b \sin\theta + T_E$$
$$N_B = sX_b \sin\theta + sY_b \cos\theta + T_N$$

(C-8)

Let $a = s \cos\theta$ and $b = s \sin\theta$. Notice that two new variables are being introduced, which are independent functions of two existing variables. This is essential so that the total number of unknown coefficients will remain the same. By substitution, Eqs. (C-8) become

$$E_A = aX_a - bY_a + T_E$$
$$N_A = aY_a + bX_a + T_N$$
$$E_B = aX_b - bY_b + T_E$$
$$N_B = aY_b + bX_b + T_N$$

(C-9)

Because both the *XY* and *EN* coordinates for points *A* and *B* are known, Eqs. (C-9) contain only four unknowns, the transformation parameters *a*, *b*, T_E, and T_N. The four equations may be solved simultaneously to obtain values for the unknowns. When the four transformation factors have been computed, an *E* and an *N* equation of the form of Eqs. (C-9) may be solved to obtain the final coordinates of each point whose coordinates were known only in the *XY* system.

By this method, the transformation can be performed without ever determining the scale and rotation parameters directly. If, for some reason, it is necessary to determine the scale and rotation parameters, these values can be derived from the values of *a* and *b* as follows:

$$s = \sqrt{a^2 + b^2} \tag{C-10}$$

$$\theta = \tan^{-1}\left(\frac{b}{a}\right) \tag{C-11}$$

In Eq. (C-11), it is necessary to use the full circle inverse tangent function (typically called 'atan2' in computer languages and spreadsheets) since the value of θ can cover the full range from –180° to +180°. With a scientific calculator, this full range can generally be achieved by using the rectangular-to-polar conversion capability.

Example C-2

Solve Example C-1 by using the alternate method.

(*a*) Formulate Eqs. (C-9) for the points whose coordinates are known in both systems.

$$1100.64 = 632.17a - 121.45b + T_E$$
$$1431.09 = 121.45a + 632.17b + T_N$$
$$1678.39 = 355.20a + 642.07b + T_E$$
$$254.15 = -642.07a + 355.20b + T_N$$

(*b*) The simultaneous solution of the above four equations yields the following:

$$a = 1.11964$$
$$b = 1.16285$$
$$T_E = 534.07$$
$$T_N = 559.99$$

(*c*) Using the four transformation parameters, the final *EN* ground coordinates of point *C* are calculated as follows:

$$E_C = 1.11964(1304.81) - 1.16285(596.37) + 534.07 = 1301.49$$
$$N_C = 1.11964(596.37) - 1.16285(1304.81) + 559.99 = 2745.01$$

(*d*) (Optional step) Compute the values for *s* and θ, using Eqs. (C-10) and (C-11), respectively.

$$s = \sqrt{(1.11964)^2 + (1.16285)^2} = 1.61425$$
$$\theta = \tan^{-1}\left(\frac{1.16285}{1.11964}\right) = 46.0845° \quad \blacktriangle$$

C-4 Coordinate Transformations With Redundancy

In some instances, more than two control points are available with coordinates known in both the

arbitrary and final systems. In that case, redundancy exists and the transformation can be computed by using a least squares solution. In this method, as discussed in App. B, the sum of the squares of the residuals in the measurements is minimized, which, according to the theory of probability, produces the most probable solution. The least squares method has the additional advantages that mistakes in the coordinates may be detected and that the precision of the transformed coordinates may be obtained. For these reasons, it is strongly advised to use redundancy in coordinate transformations whenever possible.

In the least squares procedure, it is convenient to use the alternate method discussed in Sec. C-3. Two observation equations similar to those of Eqs. (C-9) are formed for each point whose coordinates are known in both systems. Residuals v are included in the equations to make them consistent, as follows:

$$aX - bY + T_E = E + v_E$$

$$aY + bX + T_N = N + v_N \tag{C-12}$$

If n points are available whose coordinates are known in both systems, $2n$ equations may be formed containing the four unknown transformation parameters. The equations are solved by the method of least squares to obtain the most probable transformation parameters. Transformed coordinates of all required points may then be found by using the transformation factors as illustrated in step (c) of Example C-2.

It is theoretically correct in least squares to associate residuals with actual observations. In Eqs. (C-12), however, the X and Y coordinates are observed, yet residuals are only associated with the E and N control coordinates. Although there is a more rigorous least squares technique available to handle this situation, the easier approach shown above is commonly used and has been found to yield entirely satisfactory results.

C-5 Matrix Methods in Coordinate Transformations

Coordinate transformations involve rather lengthy calculations and are therefore best handled on a computer. Matrix algebra is ideal for computer calculations and is therefore convenient for performing transformations.

To illustrate the application of matrix algebra in two-dimensional conformal coordinate transformation by least squares, assume that coordinates of three control points A, B, and C are known in both the XY system and the EN system. Let their coordinates be of equal reliability so that their weights are equal. Assume also that transformation into the EN system is required for points D through N, whose coordinates are known only in the XY system.

First, six observation equations of the form of Eqs. (C-12) are developed, two for each control point A, B, and C, as follows:

$$
\begin{aligned}
X_A a - Y_A b + T_E &= E_A + v_{E_A} \\
Y_A a + X_A b + T_N &= N_A + v_{N_A} \\
X_B a - Y_B b + T_E &= E_B + v_{E_B} \\
Y_B a + X_B b + T_N &= N_B + v_{N_B} \\
X_C a - Y_C b + T_E &= E_C + v_{E_C} \\
Y_C a + X_C b + T_N &= N_C + v_{N_C}
\end{aligned}
\tag{C-13}
$$

In matrix representation, the above six equations are

$$
{}_6A^4\,{}_4X^1 = {}_6L^1 + {}_6V^1 \tag{C-14}
$$

In matrix Eq. (C-14), A is the matrix of coefficients of the unknown transformation parameters, X is the matrix of unknown transformation parameters, L is the matrix of constant terms which is made up of control point coordinates, and V is the matrix of residuals in those coordinates brought about by measurement errors. More specifically, these matrices are

$$
{}_6A^4 = \begin{bmatrix}
X_A & -Y_A & 1 & 0 \\
Y_A & X_A & 0 & 1 \\
X_B & -Y_B & 1 & 0 \\
Y_B & X_B & 0 & 1 \\
X_C & -Y_C & 1 & 0 \\
Y_C & X_C & 0 & 1
\end{bmatrix}
\qquad
{}_4X^1 = \begin{bmatrix} a \\ b \\ T_E \\ T_N \end{bmatrix}
$$

$$
{}_6L^1 = \begin{bmatrix}
E_A \\ N_A \\ E_B \\ N_B \\ E_C \\ N_C
\end{bmatrix}
\qquad
{}_6V^1 = \begin{bmatrix}
v_{E_A} \\ v_{N_A} \\ v_{E_B} \\ v_{N_B} \\ v_{E_C} \\ v_{N_C}
\end{bmatrix}
$$

As discussed in App. B, matrix Eq. (B-12) is used to solve this equally weighted system for the transformation parameters. The final transformation of all points D through N into the EN system is performed as discussed in step (*c*) of Example C-2. This phase of the computation is also readily adapted to matrix methods.

C-6 Two-Dimensional Affine Coordinate Transformation

The two-dimensional *affine* coordinate transformation is only a slight modification of the two-dimensional conformal transformation, to include anisotropic scaling (i.e., different scale factors in the x and y directions) and to compensate for nonorthogonality (nonperpendicularity or skewing) of the axis system. The affine transformation achieves these additional features by including two

additional unknown parameters for a total of six. As will be shown, the derivation of the transformation equations depends on the measurement characteristics of the arbitrary coordinate system.

A two-dimensional affine transformation consists of four basic steps: (1) scale change in x and y, (2) correction for nonorthogonality, (3) rotation, and (4) translation. Figure C-4 illustrates the geometric relationship between the arbitrary coordinate system xy and the final coordinate system XY. In this figure, the nonorthogonality of x and y is indicated by the angle ε. The rotation angle necessary to make the two systems parallel is θ, and translations T_X and T_Y account for the offset of the origin. The four steps of the derivation are as follows:

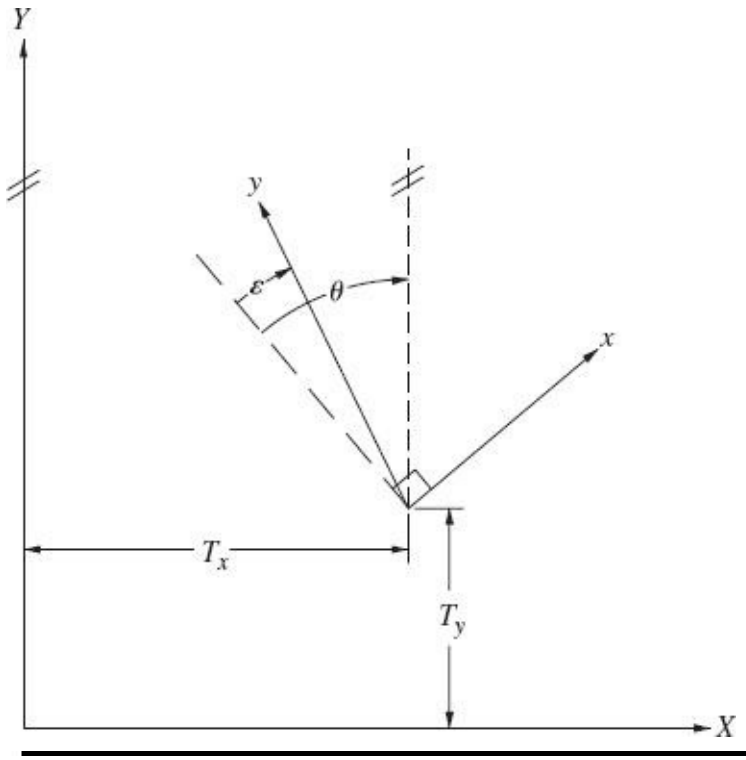


FIGURE C-4 General two-dimensional affine coordinate transformation relationship.

Step 1: Scale Change in x and y

To make the scale of the arbitrary system xy equal to that of the final system XY, each coordinate is multiplied by its associated scale factor, s_x and s_y. This results in a correctly scaled coordinate system $x'y'$ as given in Eqs. (C-15).

$$x' = s_x x$$
$$y' = s_y y \tag{C-15}$$

Step 2: Correction for Nonorthogonality

When x and y coordinates are measured from axes that intersect at a right angle, the x distance (coordinate) is measured perpendicularly from the y axis, which means the distance is at the same time parallel to the x axis. Similarly, the y distance is measured perpendicularly from the x axis and is therefore parallel to the x axis. When the xy axes are nonorthogonal, measuring an x distance perpendicularly from the y axis does not result in a distance that is parallel to the x axis. The analogous situation applies with the y distance. To derive the relationship, the specific measurement

characteristics of the system must be identified.

Several geometric possibilities exist; however, two configurations are most common and are illustrated in Figs. C-5*a* and *b*. In both figures the $x'y'$ measurement systems have already been scaled in accordance with step 1. The first configuration, illustrated in (*a*), is appropriate for most comparators where separate x and y carriages provide independent movement in both directions. The x' coordinate is measured parallel to the x' axis from the origin to the point, and the y' coordinate is measured parallel to the y' axis from the origin to the point. The second configuration, shown in (*b*), is appropriate when one is using satellite imagery that is acquired in a scanning fashion while the earth rotates beneath. The resulting image has a distinct parallelogram shape. In this configuration the x' coordinate is measured parallel to the x' axis from the y' axis to the point, and the y' coordinate is measured perpendicular to the x' axis. For the configuration of Fig. C-5*a*, the correction for nonorthogonality is given by Eqs. (C-16).

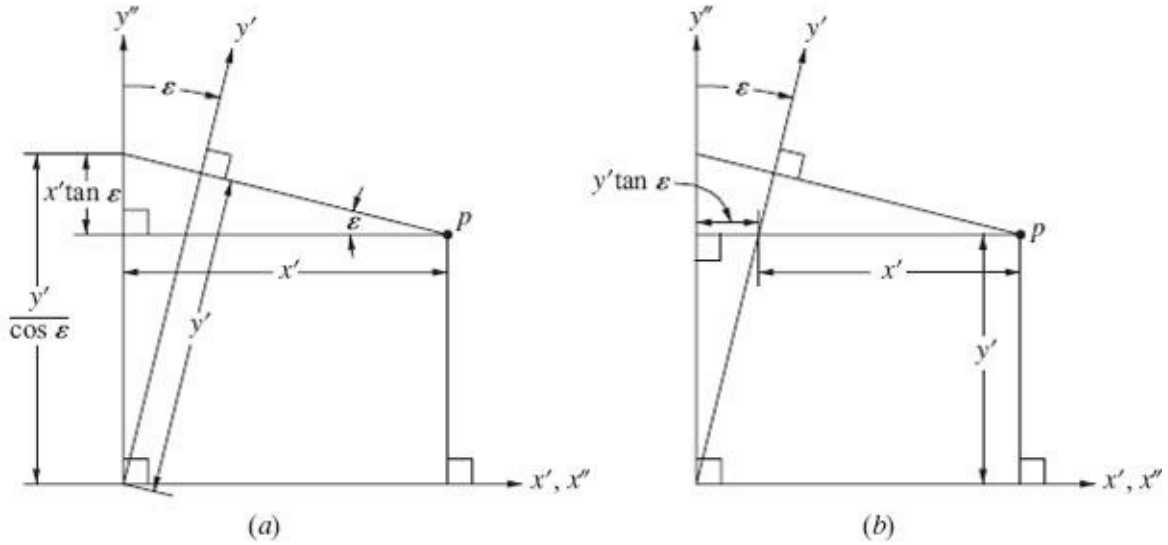


FIGURE C-5 (*a*) Two-dimensional affine relationship for typical comparator. (*b*) Two-dimensional affine relationship for typical scanning-type satellite image.

$$x'' = x'$$

$$y'' = \frac{y'}{\cos\varepsilon} - x'\tan\varepsilon \tag{C-16}$$

Equations (C-17) express the relationship for the configuration of Fig. C-5*b*.

$$x'' = x' + y'\tan\varepsilon$$

$$y'' = y' \tag{C-17}$$

Step 3: Rotation

Rotation by the angle θ is accomplished in the same fashion as in the two-dimensional conformal coordinate transformation presented in Sec. C-2. Equations (C-18) give the relationship between the $x''y''$ system and the $X'Y'$ system which is parallel to the final XY system after rotation by angle θ.

$$X' = x''\cos\theta - y''\sin\theta$$

$$Y' = x''\sin\theta + y''\cos\theta \tag{C-18}$$

Step 4: Translation

The final step is to translate the origin by T_X and T_Y to make it coincide with the origin of the final system, as shown in Eqs. (C-19).

$$X = X' + T_X$$

$$Y = Y' + T_Y \tag{C-19}$$

Combining the four steps for configuration (*a*) gives Eqs. (C-20).

$$X = s_x x\cos\theta - \left(\frac{s_y y}{\cos\varepsilon} - s_x x\tan\varepsilon\right)\sin\theta + T_X$$

$$Y = s_x x\sin\theta + \left(\frac{s_y y}{\cos\varepsilon} - s_x x\tan\varepsilon\right)\cos\theta + T_Y \tag{C-20}$$

Equations (C-20) can then be simplified as shown in the following steps, yielding Eqs. (C-21).

Step a

$$X = s_x x(\cos\theta + \tan\varepsilon\sin\theta) - s_y y\frac{\sin\theta}{\cos\varepsilon} + T_X$$

$$Y = s_x x(\sin\theta + \tan\varepsilon\cos\theta) + s_y y\frac{\cos\theta}{\cos\varepsilon} + T_Y$$

Step b

$$X = T_X + s_x x\frac{\cos\varepsilon\cos\theta + \sin\varepsilon\sin\theta}{\cos\varepsilon} - s_y y\frac{\sin\theta}{\cos\varepsilon}$$

$$Y = T_Y + s_x x\frac{\cos\varepsilon\sin\theta + \sin\varepsilon\cos\theta}{\cos\varepsilon} + s_y y\frac{\cos\theta}{\cos\varepsilon}$$

Step c

$$X = T_X + s_x x \frac{\cos(\varepsilon - \theta)}{\cos\varepsilon} - s_y y \frac{\sin\theta}{\cos\varepsilon}$$

$$Y = T_Y - s_x x \frac{\sin(\varepsilon - \theta)}{\cos\varepsilon} + s_y y \frac{\cos\theta}{\cos\varepsilon}$$

(C-21)

To simplify solutions involving Eqs. (C-21), the following substitutions are made.

Let

image

Making these substitutions into Eqs. (C-21) gives Eqs. (C-22), which are the final form of the affine transformation. In Eqs. (C-22), the six unknown parameters s_x, s_y, ε, θ, T_X, and T_Y, which appeared in Eqs. (C-21) in a nonlinear form, have been replaced by six independent parameters a_0, a_1, a_2, b_0, b_1, and b_2, resulting in a linear form.

(C-22)

After solving an affine transformation using Eqs. (C-22), if it is necessary to obtain values for the original six parameters for configuration (*a*), they may be obtained as follows. Note that in the first two expressions, the full circle inverse tangent function (e.g., atan2) must be used.

image

Combining the four steps for configuration (*b*) gives Eqs. (C-23).

image

(C-23)

Equations (C-23) can then be simplified as shown in the following steps, yielding Eqs. (C-24).

Step a

Step b

image

Step c

image

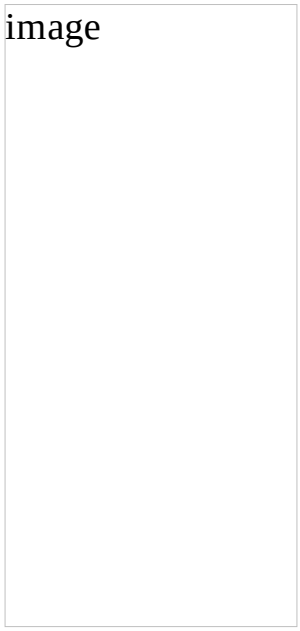
(C-24)

To simplify solutions involving Eqs. (C-24), the following substitutions are made.

Let

image

After making these substitutions into Eqs. (C-24), the same linear form of the affine transformation is obtained as before, i.e., Eqs. (C-22). After solving an affine transformation using Eqs. (C-22), if it is necessary to obtain values for the original six parameters for configuration (*b*), they may be obtained as follows. Note that in the first two expressions, the full circle inverse tangent function (e.g., atan2) must be used.

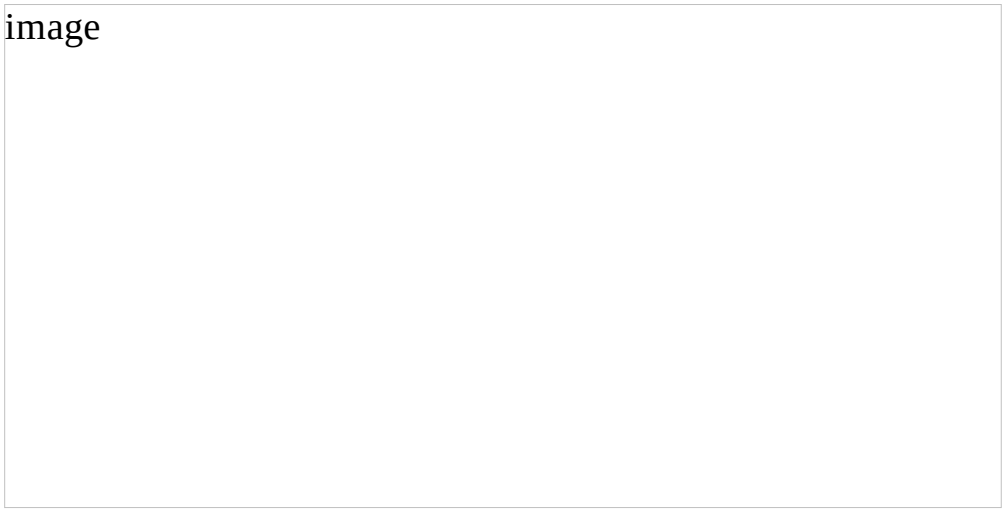


As noted above, the transformation equations [Eqs. (C-22)] are identical for either configuration (*a*) or (*b*), and thus the method of solution is the same for both. The two derivations have been presented, however, to enable computation of the transformation parameters after solution, a factor that can be important in certain situations, for example in evaluating or calibrating equipment.

As with the two-dimensional conformal transformation, the application of the affine transformation is a two-step procedure of: (1) determining the *a* and *b* coefficients, using points whose coordinates are known in both the *XY* and *xy* systems, and (2) applying these coefficients to calculate transformed *XY* coordinates for all other points from their *xy* coordinates. In correcting photo coordinates in film photography, the fiducial marks are used to perform step 1, since their calibrated *XY* coordinates are known from camera calibration, and their *xy* coordinates are available from comparator measurements. For a given photograph, a pair of equations of the form of Eqs. (C-22) can be written for each fiducial mark. If there are four fiducials, four *X* and four *Y* equations are obtained. Any three of the fiducial marks could be used to obtain a solution for the unknown *a*'s and *b*'s. An improved solution may be obtained, however, if all four fiducial marks are used and the system is solved by least squares. In other applications where the geometric arrangement of common points (control) is not predetermined, it is important that the control points do not approximate a straight line, since this would lead to a very weak or even an indeterminate solution.

Example C-3

Calibrated coordinates and comparator-measured coordinates of the four fiducial marks for a certain photograph are given in the following table. The comparator-measured coordinates of other points 1, 2, and 3 are also given. It is required to compute the corrected coordinates of points 1, 2, and 3 by using the affine transformation.

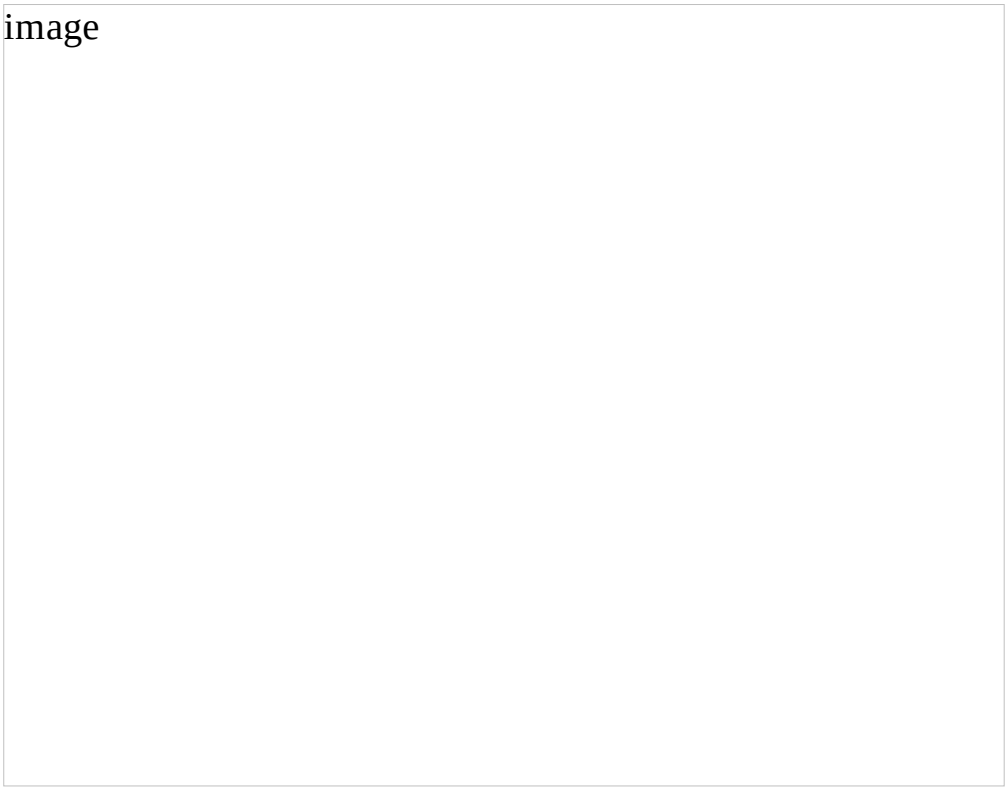


Solution Equations of the form of Eqs. (C-22), with residuals added for consistency, are formulated for the four fiducial marks as follows:

In matrix representation, after switching of the terms to the opposite side, the above eight equations are

(C-25)

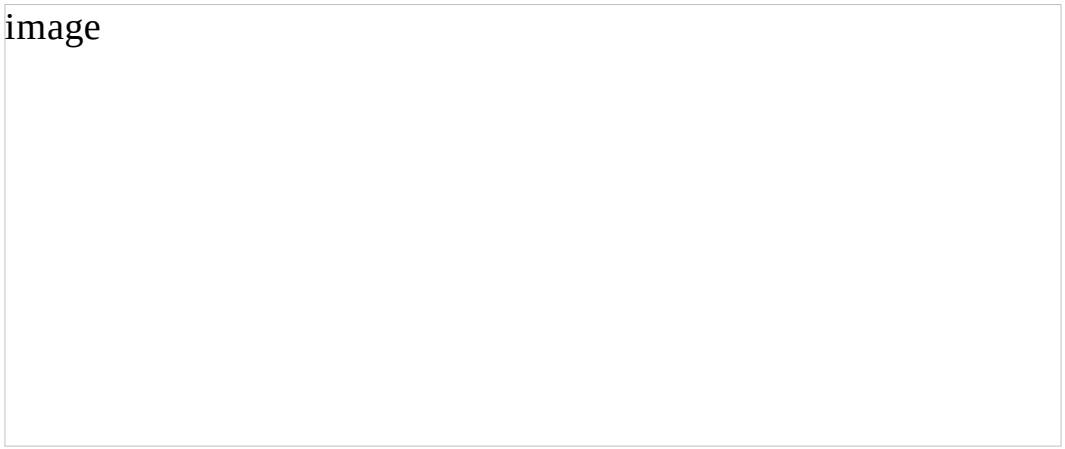
In matrix Eq. (C-25), A is the matrix of coefficients of the unknown transformation parameters, X is the matrix of unknown transformation parameters, L is the matrix of constant terms which consists of calibrated fiducial coordinates, and V is the matrix of residuals in those coordinates brought about by measurement errors. More specifically, these matrices are



Solving this system of equations using the method of least squares with equal weights as presented in Sec. B-9, the following solution is obtained:

Calculation of residual matrix *V* gives the following:

Since all the residuals have small magnitudes, the solution is deemed acceptable and the transformed coordinates of points 1, 2, and 3 are then computed as follows:



In computing affine transformations, it is strongly recommended that more than the minimum number of common points be used and that the solution be obtained by least squares. In addition to other benefits, least squares can reveal the presence of measurement mistakes through calculation of residuals.

C-7 Three-Dimensional Conformal Coordinate Transformation

As implied by its name, a three-dimensional conformal coordinate transformation involves converting from one three-dimensional system to another. In the transformation, true shape is retained. This type of coordinate transformation is essential in analytical or computational photogrammetry for two basic problems: (1) to transform arbitrary stereomodel coordinates to a ground or object space system and (2) to form continuous three-dimensional "strip models" from independent stereomodels. Three-dimensional conformal coordinate transformation equations are developed here in general, while their application to specific photogrammetry problems is described elsewhere in the text where appropriate.

In Fig. C-6, it is required to transform coordinates of points from an *xyz* system to an *XYZ* system. As illustrated in the figure, the two coordinate systems are at different scales, they are not parallel, and their origins do not coincide. The necessary transformation equations can be expressed in terms of seven independent parameters: the three rotation angles omega (ω), phi (φ), and kappa (κ); a scale factor s; and three translation parameters T_X, T_Y, and T_Z. Before proceeding with the development of the transformation equations, it is important to define sign conventions. All coordinates shall be defined as righthanded, i.e., systems in which positive X, Y, and Z are as shown in Fig. C-6. Rotation angles ω, φ, and κ are positive if they are counterclockwise when viewed from the positive end of their respective axes. For example, positive ω rotation about the x' axis, is shown in Fig. C-6.

FIGURE C-6 *XYZ* and *xyz* right-handed three-dimensional coordinate systems.

The transformation equations will be developed in the following two basic steps: (1) rotation and (2) scaling and translation.

Step 1: Rotation

In Fig. C-6, an $x'y'z'$ coordinate system parallel to the *XYZ* object system is constructed with its origin at the origin of the *xyz* system. In the development of rotation formulas, it is customary to consider the three rotations as taking place so as to convert from the $x'y'z'$ system to the *xyz* system. The rotation equations are developed in a sequence of three independent two-dimensional rotations. These rotations, illustrated in Fig. C-7, are, first, ω rotation about the x' axis which converts coordinates from the $x'y'z'$ system to an $x_1y_1z_1$ system; second, ϕ rotation about the once-rotated y_1 axis which converts coordinates from the $x_1y_1z_1$ system to an $x_2y_2z_2$ system; and third, κ rotation about the twice-rotated z_2 axis which converts coordinates from the $x_2y_2z_2$ system to the *xyz* system of Fig. C-6. The exact amount and direction of the rotations for any three-dimensional coordinate transformation will depend upon the orientation relationship between the *xyz* and *XYZ* coordinate systems.

FIGURE C-7 The three sequential angular rotations.

The development of the rotation formulas is as follows:

1. Rotation through the angle ω about the x' axis is illustrated in Fig. C-8. The coordinates of any point A in the once-rotated $x_1y_1z_1$ system, as shown graphically in Fig. C-8, are

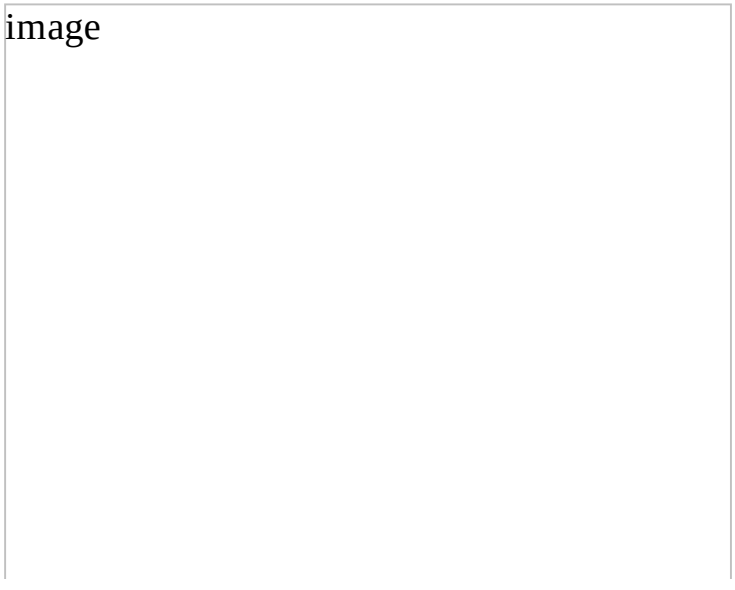


FIGURE C-8 Omega rotation about the x' axis.

image

(C-26)

Since this rotation was about x', the x' and x_1 axes are coincident and therefore the x coordinate of A is unchanged.

2. Rotation through ϕ about the y_1 axis is illustrated in Fig. C-9. The coordinates of A in the twice-rotated $x_2y_2z_2$ coordinate system, as shown graphically in Fig. C-9, are

image

FIGURE C-9 Phi rotation about the y_1 axis.

image

(C-27)

In this rotation about y_1, the y_1 and y_2 axes are coincident, and therefore the y coordinate of A is unchanged. Substituting Eqs. (C-26) into (C-27) gives

image

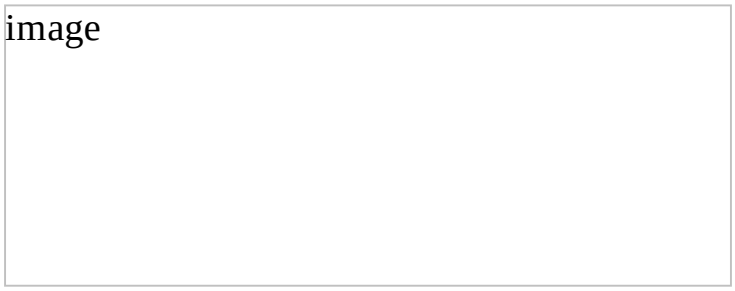
(C-28)

3. Rotation through κ about the z_2 axis is illustrated in Fig. C-10. The coordinates of A in the three-times-rotated coordinate system, which has now become the xyz system as shown graphically in Fig. C-10, are

FIGURE C-10 Kappa rotation about the z_2 axis.

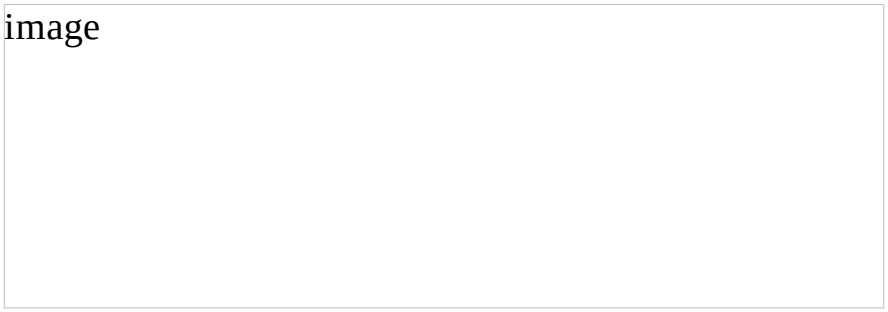
(C-29)

In this rotation about z_2, the z_2 and z axes are coincident, and therefore the z coordinate of A is unchanged. Substituting Eqs. (C-28) into (C-29) gives



(C-30)

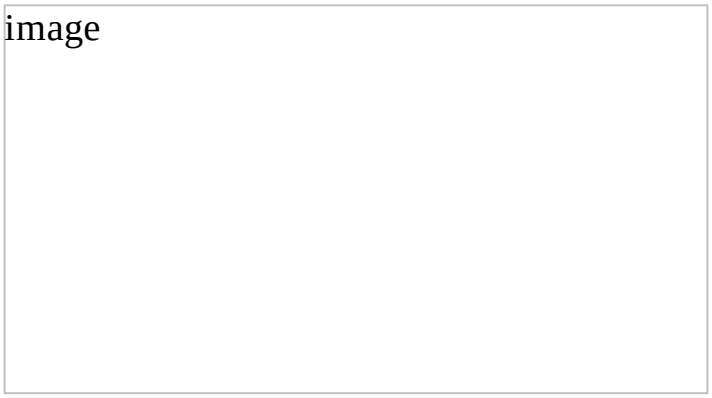
Factoring Eqs. (C-30) gives



(C-31)

Substituting *m's* for the coefficients of *x'*, *y'*, and *z'* in Eqs. (C-31) gives

(C-32)



(C-33)

Equations (C-32) may be expressed in matrix form as

(C-34)

where

The matrix *M* is commonly called the *rotation matrix*. The individual elements of the rotation matrix are *direction cosines* which relate the two axis systems. These matrix elements, expressed in terms of direction cosines, are

(C-35)

In the above matrix, cos xx' is the direction cosine relating the x and x' axes, cos xy' relates the x and y' axes, etc. Direction cosines are simply the cosines of the angles in space between the respective axes, the angles being taken between 0° and 180°. It is an important property that *the sum of the squares of the three direction cosines in any row or in any column is unity*. This property may be used to check the computed elements of the rotation matrix for correctness.

The rotation matrix is an *orthogonal* matrix, which has the property that *its inverse is equal to its transpose*, or

(C-36)

By using this property, Eq. (C-34) may be rewritten, expressing $x'y'z'$ coordinates in terms of xyz coordinates as follows:

(C-37)

In expanded form, this equation is

(C-38)

Step 2: Scaling and Translation

To arrive at the final three-dimensional coordinate transformation equations, i.e., equations that yield coordinates in the XYZ system of Fig. C-6, it is necessary to multiply each of Eqs. (C-38) by a scale factor s and to add the translation factors T_X, T_Y, and T_Z. [Recall that the $x'y'z'$ coordinates given by Eqs. (C-38) are in a system that is parallel the XYZ system.] This step makes the lengths of any lines equal in both coordinate systems, and it translates from the origin of $x'y'z'$ to the origin of the XYZ system. Performing this step yields

(C-39)

In matrix form, Eqs. (C-39) are

(C-40)

In Eq. (C-40), matrices M and X are as previously defined, s is the scale factor, and

In Eqs. (C-39), the nine *m's* are not independent of one another, but rather, as seen in Eqs. (C-33), they are functions of the three rotation angles ω, ϕ, and κ. In addition to these three unknown angles, there are three unknown translations and one scale factor in Eqs. (C-39), for a total of seven unknowns. A unique solution is obtained for the unknowns if the x and y coordinates of two horizontal points and the z coordinates of three vertical points are known in both coordinate systems. If more than the minimum of seven coordinates are known in both systems, redundant equations may be written, which makes possible an improved solution through least squares techniques.

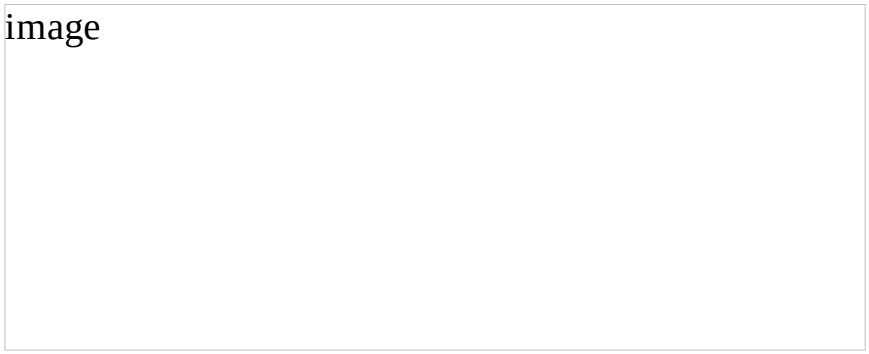
Solution of the three-dimensional conformal coordinate transformation is more complex than that of the two-dimensional transformations presented earlier in this appendix. This added complexity is due to the fact that Eqs. (C-39) are nonlinear in terms of the unknowns s, ω, ϕ, and κ. While it is possible to directly compute values for the translation factors since their terms exist in a linear form, it is more convenient to treat them as if they appeared in a nonlinear form. To solve these equations, they are linearized by using a Taylor series expansion including only the first-order terms. Application of Taylor's series requires that initial approximations be obtained for each of the seven unknowns. Each point P contributes as many as three equations (two if the point is horizontal only and one if the point has just vertical control) which are linearized as shown in Eqs. (C-41).

image

(C-41)

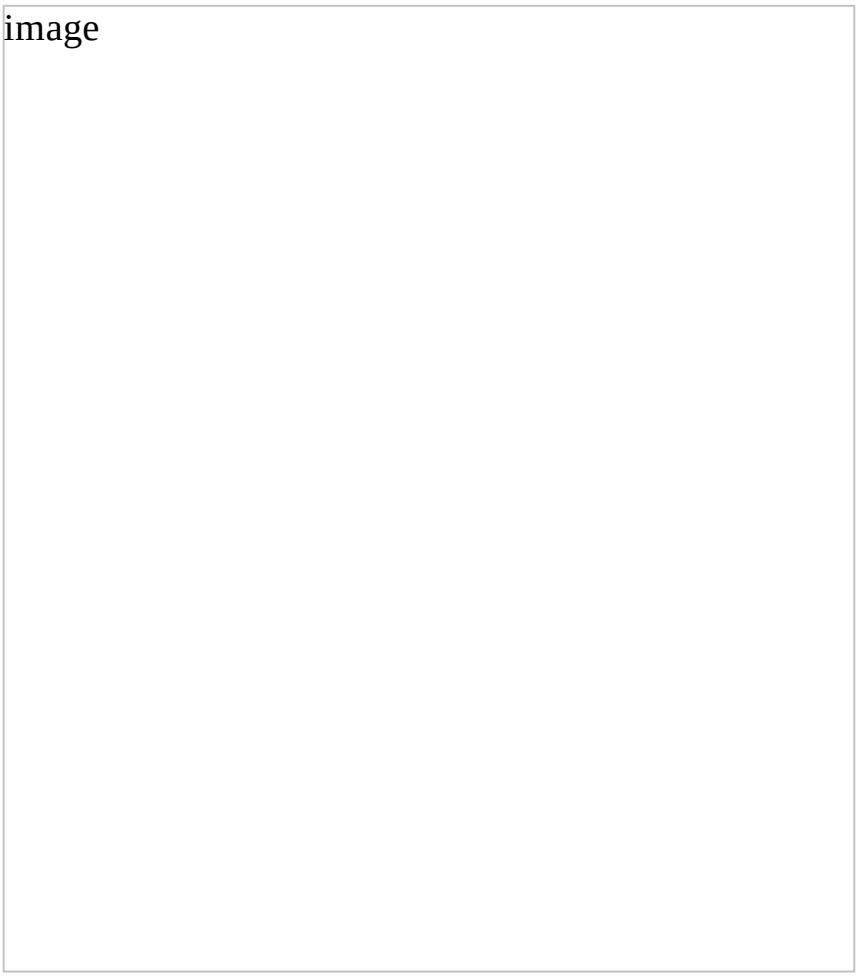
In Eqs. (C-41), $(X_P)_0$, $(Y_P)_0$, and $(Z_P)_0$ are the right-hand sides of Eqs. (C-39) evaluated at the initial approximations; $(\partial X_P/\partial s)_0$, $(\partial X_P/\partial \omega)_0$, etc. are the partial derivatives with respect to the indicated unknowns evaluated at the initial approximations; and ds, $d\omega$, $d\phi$, $d\kappa$, dT_X, dT_Y, and dT_Z are corrections to the initial approximations which will be computed during the solution. The units of $d\omega$, $d\phi$, and $d\kappa$ are radians.

Substituting letters for partial derivative coefficients, adding residuals to make the equations suitable for a least squares solution, and rearranging terms, the following equations result.



(C-42)

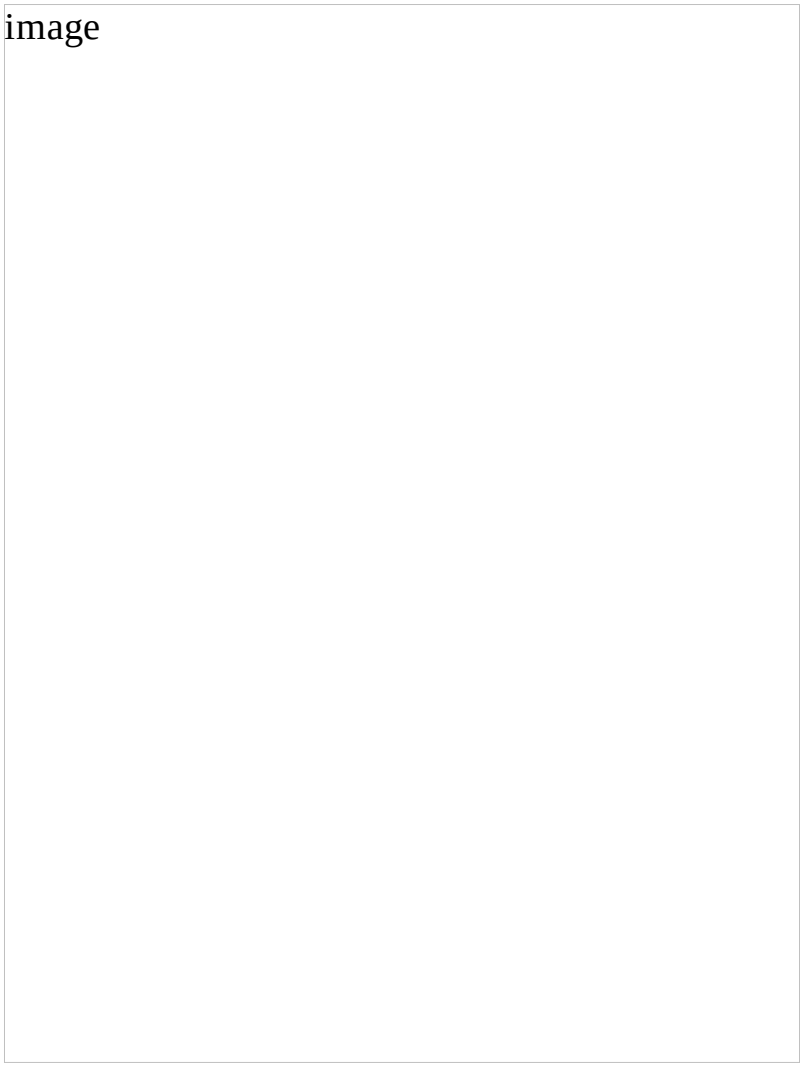
To clarify the coefficients of Eqs. (C-42), the partial derivative terms, which must be evaluated at the initial approximations, are as follows:



In a least squares solution, each point contributes up to three rows of coefficients to the *A* matrix, as well as terms in the *L* and *V* matrices. In general, assuming points *1*, *2*, …, *n* are three-dimensional control points, the following matrix equation results:

(C-43)

where



Equation (C-43) may be solved by using least squares Eq. (B-11), giving corrections which are then added to the initial approximations for the unknowns, resulting in a better set of approximations. The solution must be iterated (since only the first-order terms of Taylor's series were used) until negligibly small values are obtained in matrix *X*. Other techniques of testing for the convergence of an iterative solution, such as convergence of the computed estimate of the standard deviation of unit weight [see Eq. (B-15) for an unweighted adjustment or Eq. (B-16) for a weighted adjustment], may also be used. The reader may consult references listed at the end of this appendix for a discussion of these convergence techniques.

Once the solution has reached satisfactory convergence, the latest approximations for the unknowns are the values for the transformation parameters. Then the transformed coordinates for each point whose coordinates are known only in the original system are obtained by applying Eqs. (C-39).

C-8 Initial Approximations for the 3D Conformal Coordinate Transformation

In solving nonlinear equations using Taylor's theorem, it is important that good initial approximations be obtained. If approximations for the 3D conformal coordinate transformation parameters are not reasonably close to the final computed values, the solution will generally diverge (i.e., the corrections get larger with each subsequent iteration instead of smaller), or the solution may converge to an erroneous set of transformation parameters. In cases involving stereomodels formed from near-

vertical photography, ω and ϕ may be assumed to be zero. An approximation for κ may be determined from the difference in azimuths of a common line in both systems, and an approximation for s may be determined from the ratio of the length of a line in the control system over the length of the same line in the arbitrary system. Approximations for the translation factors may then be found by rearranging Eqs. (C-39), so that T_X, T_Y, and T_Z are isolated on the left side. By using previously determined initial approximations for s, ω, ϕ and κ, and the two sets of coordinates for a common point on the right-hand side, initial approximations for the translations may thus be determined.

When the assumption of near-vertical photography is invalid, such as when one is dealing with terrestrial photography, an alternative method must be used to approximate ω, ϕ, and κ. One method for accurately approximating these values is presented here. The method uses the azimuth-tilt-swing (α-t-s) conversion to ω-ϕ-κ. A description of rotation in terms of azimuth, tilt, and swing can be found in Sec. D-9 and the method for conversion from α-t-s to ω-ϕ-κ can be found in Sec. D-11. The approximation method can be divided into the following five steps:

Step 1: Compute the Normal Vectors for Three Points

The first step is to calculate the normal vectors to the plane formed by three points in both the arbitrary and control systems. The normal vector, n, is defined as the vector orthogonal to the plane as illustrated in Fig. C-11*a*. It can be calculated by using Eq. (C-44), the cross product of the vector from point 1 to point 2 (p_{12}) and the vector from point 1 to point 3 (p_{13}). The components of the normal vector relative to the origin are illustrated in Fig. C-11*b*.

image

(C-44)

image

FIGURE C-11 (*a*) Normal vector, **n**, relative to the plane defined by three points. (*b*) Normal vector, **n**, relative to the origin.

For the best results, the three points used to calculate the normal vectors should be chosen so that they have the strongest geometric configuration of all possible combinations of three points. This can be achieved by choosing three points that are the least collinear. The points that are the least collinear are those that form a triangle with the largest *altitude*, *h*, the perpendicular distance from the longest

side to the point not on that side. This is illustrated in Fig. C-12. The formula for the square of the altitude—where a is the length of the longest side of the triangle—is shown in Eq. (C-45).

(C-45)

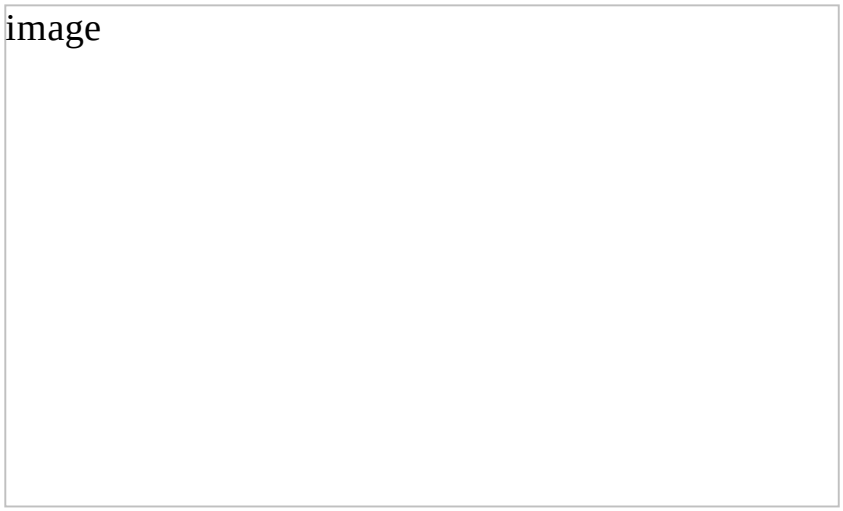


FIGURE C-12 Altitude, h, of the triangle formed by three points.

Step 2: Calculate Tilt and Azimuth Using the Normal Vectors

The tilt of the normal vectors from both the arbitrary and control coordinates is the magnitude of their angle from vertical, i.e., the negative z axis. The azimuth of the normal vector is the azimuth of the projection of it onto the x-y plane. Values for these angles can be calculated using Eqs. (C-46) and (C-47) from the components of the normal vector illustrated in Fig. C-11*b*. Note that the full circle inverse tangent must be used in Eq. (C-47) to obtain the proper quadrant.

(C-46)

(C-47)

Step 3: Rotate the Points in Both Systems

Rotation matrices are formed, where the tilts and azimuths are those calculated in step 2 and the swings are set to zero, for both the arbitrary and control coordinate systems. Equations (D-28) can be used to do this. These matrices can then be applied to two points in both systems, resulting in the lines between the points being horizontal, i.e., having equal z values.

Step 4: Calculate Swing for the Common Line

The azimuths of the rotated lines are calculated using the rotated points from step 3. The difference in

azimuth of these lines is the swing required to align the rotated arbitrary system with the rotated control system. This intermediate swing is calculated using Eq. (C-48).

(C-48)

Step 5: Combine the Two Tilts, Two Azimuths, and One Swing into a Single Rotation Matrix to Obtain Omega, Phi, and Kappa

A rotation matrix, M_a, is formed using the tilt and azimuth from step 3 for the arbitrary system and the swing found in step 4. Likewise, the tilt and azimuth (swing is set to zero) for the control system found in step 3 are used to form M_c. To form the overall rotation matrix, we multiply M_c by the transpose of M_a as shown in Eq. (C-49).

(C-49)

The rotation parameters omega, phi, and kappa can be then obtained from M using the method described in Sec. D-10.

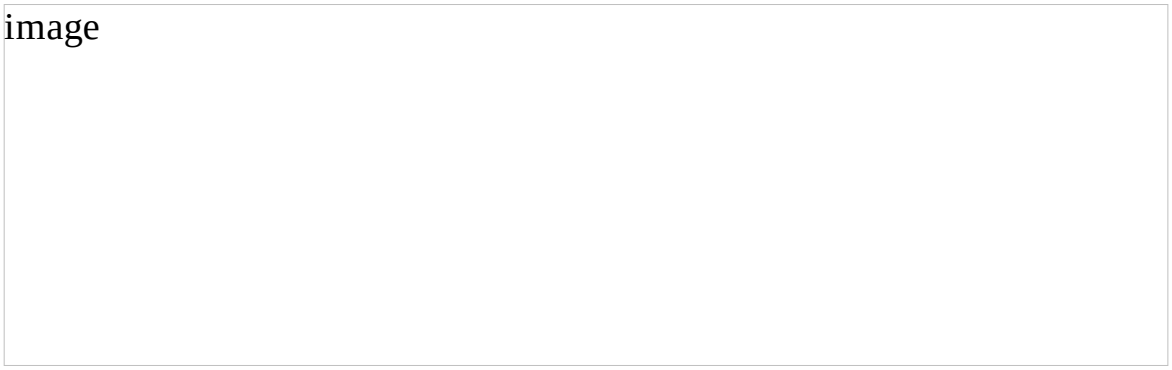
Example C-4

Initial approximations are needed to solve for the 3D conformal coordinate transformation from a close-range stereomodel to an object space system. Find initial approximations for all seven transformation parameters given the following control and arbitrary coordinates:

image

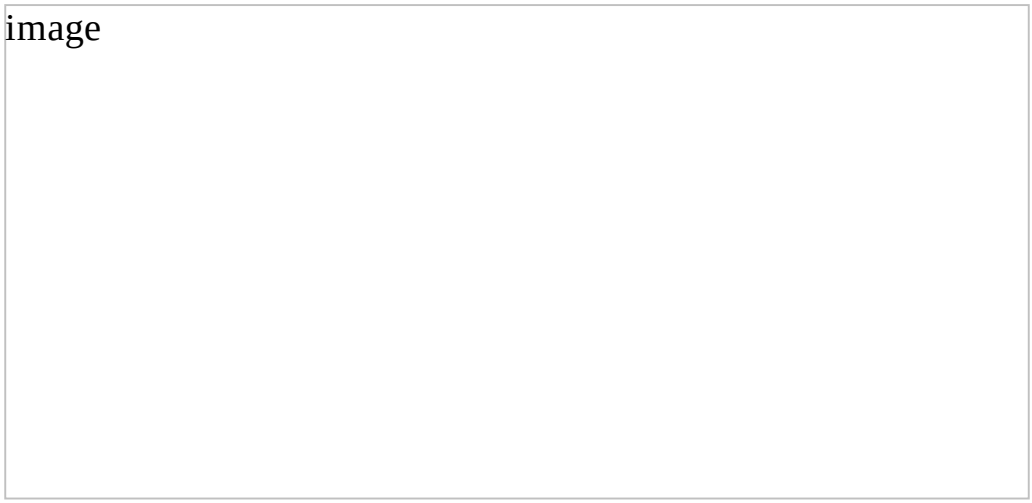
Solution The solution for the transformation parameters can be found by using the following steps:

Step 1: First, Eq. (C-45) is used to find the best geometrically configured points. The points that form the triangle with the largest altitude are 101, 102, and 104, with h = 15.303 m in control coordinates. Thus, these points are selected for calculation of the initial approximations for the transformation parameters. The normal vector for the arbitrary coordinates, $\mathbf{n}_a$ is found using Eq. (C-44).



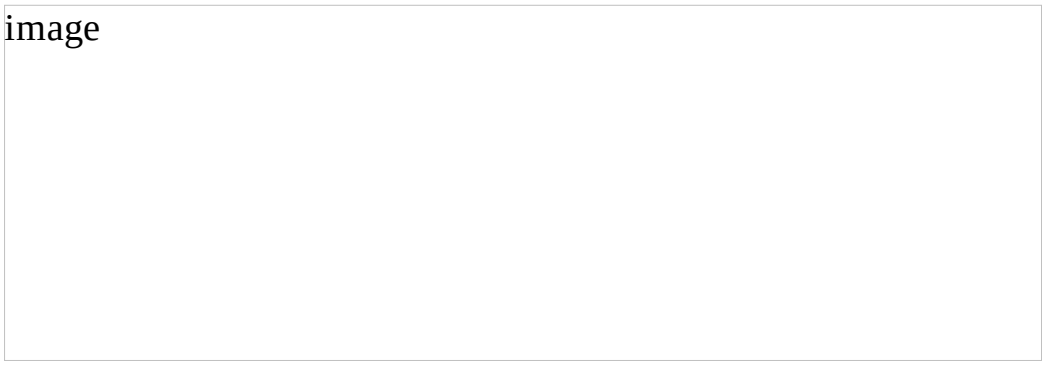
Similarly, the normal vector for the control coordinates, $\mathbf{n}_c$, is the following:

Step 2: The tilt and azimuth for each normal vector are calculated using Eqs. (C-46) and (C-47):



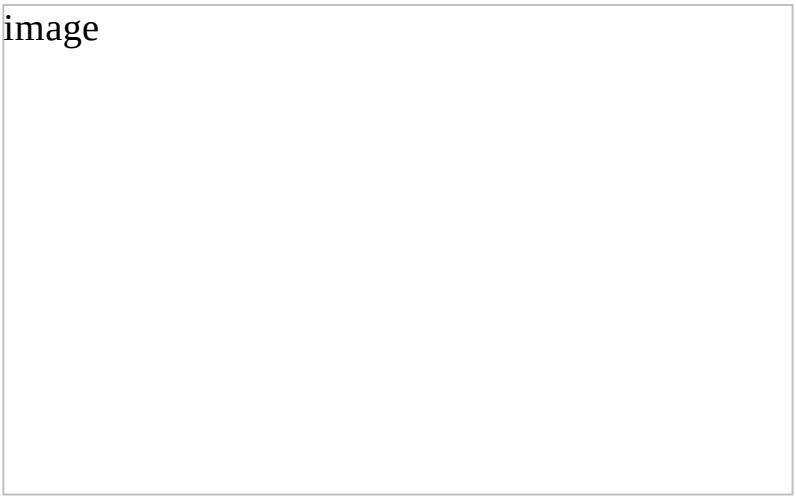
Step 3: Two rotation matrices are formed using the method in Eqs. D-28 and the tilts and azimuths found in step 2 (swing is set to zero). The matrices are then applied to points 101 and 102 in both systems yielding the coordinates of two pairs of points horizontal to the *x*-*y* plane. Points 101 and 102 in the arbitrary system:

Points 101 and 102 in the control system:



Step 4: Swing between the rotated lines, the difference in azimuth between the rotated points from step 3, is calculated using Eq. (C-48).

Step 5: Rotation matrix M_c is the same as the one used on the control coordinates in step 3. M_a is formed using tilt_a and azimuth_a in step 2 and swing calculated in step 4. These are used to calculate the overall rotation matrix.



The method described in Sec. D-11 is then used to obtain ω, ϕ, and κ.

An approximation for scale can be obtained by taking the average ratio of distances between each

pair of common points in both coordinate systems, which results in $s = 0.2997$. Approximations for translations can be found by taking the average of

using all common points, which results in $T_X = -620{,}450.008$ m, $T_Y = 96{,}100.010$ m, and $T_z = 1.185$ m.

Using these initial approximations, the first iteration corrections to the transformation parameters are

image

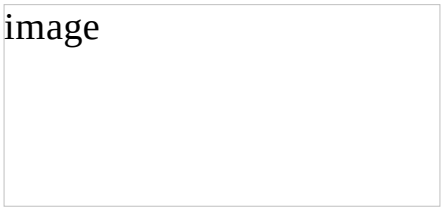
These small corrections indicate that the initial approximations were very close to the final value.

C-9 Two-Dimensional Projective Coordinate Transformation

Two-dimensional projective transformation equations enable the analytical computation of the *XY* coordinates of points after they have been projected onto a plane from another nonparallel plane. The most common use of these equations is in analytical rectification, i.e., calculating coordinates of points in a rectified-ratioed photo plane based upon their coordinates in a tilted photograph. This situation is illustrated in Fig. C-13. In the figure, a tilted photo with its *xy* photo coordinate axis system (shown in the plane of the photo) is illustrated. The photo coordinates of the principal point *o* are $x_o y_o$. The projection center (the origin of the *x'y'z'* system) is at *L*, which is on the end of a perpendicular to the photo plane from point *o*, such as $Lo = f$. The *x'y'* axes are parallel to the photo *xy* axes. The projection of points *a*, *b*, *c*, and *d* from the tilted photo onto the plane of the rectified-ratioed photo occurs at *A*, *B*, *C*, and *D*, respectively. Positions of projected points in the rectified-ratioed photo plane are expressed in the *XYZ* coordinate system shown in the figure.

In a simplified development of the equations for the two-dimensional projective transformation, an *X'Y'Z'* coordinate system is adopted which is parallel to the *XYZ* system and has its origin at *L*. Using Eqs. (C-32), which were developed in the preceding section, the *x'y'z'* coordinates of any point, such as *p* of Fig. C-13, can be expressed in terms of *X'Y'Z'* coordinates as follows:

FIGURE C-13 Geometry of two-dimensional projective transformation.



(C-50)

In Eqs. (C-50), the *m's* are functions of the rotation angles omega, phi, and kappa which define the tilt relationships between the two plane coordinate systems *xy* and *XY*. These functions are described in the preceding section. The other terms in Eqs. (C-50) are coordinates as previously described. Consider now Fig. C-14, which shows the parallel relationships between the *X'Y'Z'* and *XYZ* planes after rotation. From similar triangles of Fig. C-14,

FIGURE C-14 Parallel relationships that exist after rotation between $X'Y'Z'$ and XYZ planes in two-dimensional projective transformation.

from which

(*a*)

Again from similar triangles of Fig. C-14,

from which

(*b*)

Also, intuitively the following equation may be written:

(*c*)

Substituting Eqs. (*a*), (*b*), and (*c*) into Eq. (C-50) gives

image

(C-51)

Factoring Eqs. (C-51) gives

(*d*)

(*e*)

(*f*)

Dividing Eqs. (*d*) and (*e*) by Eq. (*f*) yields

(*g*)

(*h*)

Referring to Fig. C-13, it can be seen that the $x'y'$ coordinates are offset from the photo coordinates xy by x_o and y_o. Furthermore, for a photograph, the z coordinates are equal to $-f$. The following equations provide for these relationships:

(*i*)

(*j*)

(*k*)

Substituting Eq. (*k*) into Eqs. (*g*) and *h*) and rearranging gives

(*l*)

(*m*)

Equations (*l*) and (*m*) contain a total of seven unknown parameters (f, ω, ϕ, κ, X_L, Y_L, and Z_L) as well as the measured coordinates , , X_p, and Y_p in the two coordinate systems. Two of the unknown parameters, f and Z_L, however, are not independent. By using the ratio of these two dependant parameters as a constant, a single parameter, H can be used in their place. The relationship is shown in Eq. (*n*).

(*n*)

Given Eq. (*n*), substitutions can be made for f and Z_L in Eqs. (*l*) and (*m*) such that the ratio of the two parameters is maintained. A value of 1 (unity) will be used in place of f, and H will be used in place of Z_L, giving the following equations:

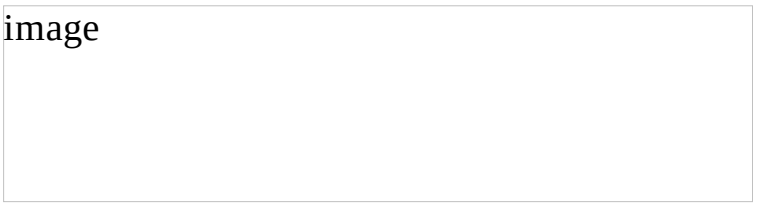
image

(C-52)

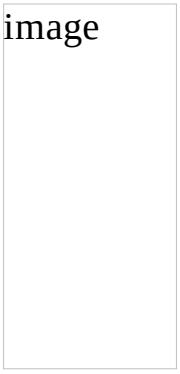
Equation (C-52) in its present form provides the *perspective projection* of a plane onto a tilted plane. Perspective projection is the true physical relationship between an image coordinate system and an object coordinate system. However, it is sometimes advantageous to generalize the relationship in Eq. (C-52) to simplify calculations. The subscript *p's* will be dropped and and will be replaced with x and y at the start of this process. First, separate the terms containing X and Y in both the numerator and the denominator.

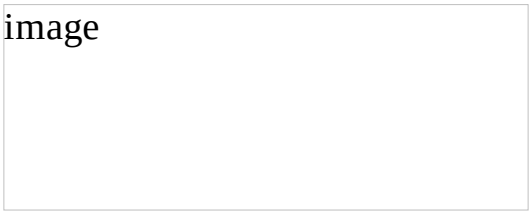
Next, use the following definition:

and multiply both the numerator and the denominator by s.



Finally, we can define the following transformation parameters:





This gives the form of the two-dimensional projective transformation equations.

(C-53)

Notice that in Eqs. (C-53) there are eight transformation parameters, functions of the six independent transformation parameters ω, ϕ, κ, X_L, Y_L, and H, from Eq. (C-52). In order to maintain a

perspective projection relationship, one must reduce the degrees of freedom back to six by enforcing two constraints on the eight transformation parameters of Eqs. (C-53). The constraints are based on the orthogonality of rotation matrices described by Eq. (C-36).

image

However, these constraints are nonlinear, and are therefore not typically used since Eqs. (C-53) can be rearranged into a linear form as will be shown later in this section. It should be noted that if the constraints are not used, Eqs. (C-53) can result in coordinate transformations that are not possible with true *perspective projection*, specifically the anisotropic scaling and skewing characteristic of affine transformations. Thus it is simply called a *projective* transformation.

As developed, Eqs. (C-53) yield x and y tilted photo coordinates from X and Y rectified-ratioed coordinates. It is customary, however, to perform rectification in the opposite sense, i.e., to compute X and Y rectified-ratioed coordinates from x and y coordinates measured on a tilted photo. Since Eqs. (C-53) are general, they can be written in the following form to enable computation of X and Y rectified-ratioed coordinates in terms of x and y tilted photo coordinates.

image

(C-54)

In using Eqs. (C-54) for rectification, X and Y are ground coordinates of control points, and x and y are coordinates of the same points in the photo coordinate system of the tilted photograph. As listed, these equations are nonlinear due to their rational form. To facilitate their use with least squares, Eqs. (C-54) can be rearranged as follows:

image

(C-55)

In Eqs. (C-55), measured xy and XY coordinates appear on the left-hand side. Technically, equations of this form, as with all transformation equations in this section, should be treated by general least squares techniques instead of the simpler method of observation equations. If, however, the measured values of x, y, X, and Y on the left side of the equations are treated as constants and if X and Y on the right side are treated as measured values, results sufficient for initial approximations, and in some cases, final values can be obtained.

A pair of Eqs. (C-55) can be written for each control point, and since there are eight unknown parameters, four control points are needed for a unique solution. It is strongly recommended that more than four control points be used in the least squares solution, however. When equations of the type of Eqs. (C-55) have been written for all control points, a solution can be obtained for the unknown parameters.

These parameters can be used as final results to compute the rectified and ratioed coordinates of all other points whose xy coordinates have been measured, or they can serve as initial approximations

for a nonlinear least squares solution to Eqs. (C-54).

Once initial approximations are obtained for the eight transformation parameters, the linearized form of Eqs. (C-54) can be used to find the least squares solution. Equations (C-56) show the linearized form of Eqs. (C-54), and Eqs. (C-57) shows the least squares observation equations. Note that in the linearized equations, the partial derivatives with respect to parameters that do not occur in the equations are omitted since their value is equal to zero (e.g., $(\partial Y_P/\partial a_1)_0 = 0$), however, the subscripts of the partial derivative terms in Eqs. (C-57) are kept consistent to facilitate the parameterization of the *A* matrix.

image

(C-56)

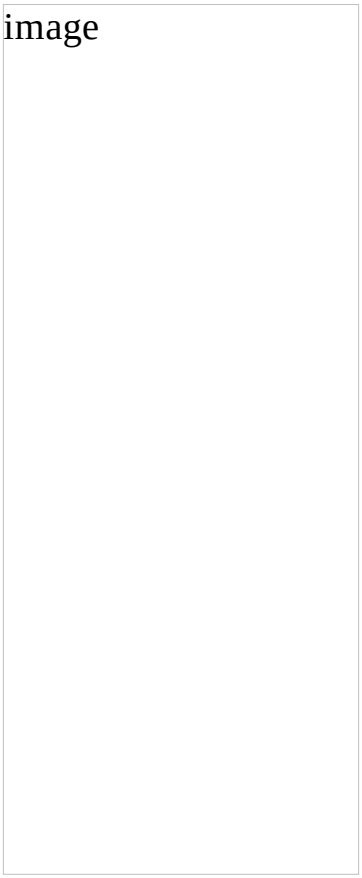
image

(C-57)

In the least squares solution, each point contributes two rows of coefficients to the *A*, *L*, and *V* matrices. In a similar manner as the three-dimensional conformal coordinate transformation, for *n* points the matrix equation in Eq. (C-58) is used. The solution is iterated until corrections in matrix *X* are negligibly small.

(C-58)

The coefficients of Eqs. (C-57) evaluated at the initial approximations are as follows:



Besides using the two-dimensional projective transformation for rectification, it can be used to transform comparator coordinates into the photo coordinate system defined by fiducial marks in film photography. However, this should only be done if more than four fiducials are available. In this instance, X and Y are calibrated coordinates of fiducial marks, and x and y are their comparator coordinates.

C-10 Two-Dimensional Rational Polynomial Coordinate Transformation

Many coordinate transformations have a similar formulation. For example, the two-dimensional affine coordinate transformation in Sec. C-6 can be considered a specific form of the two-dimensional projective coordinate transformation in Sec. C-9. That is, if $a_3 = b_3 = 0$ in Eqs. (C-54), the result is effectively an affine relationship. Continuing with this idea, we can formulate a general equation for some two-dimensional coordinate transformations as a rational polynomial shown in Eqs. (C-59).

(C-59)

It is important when using high-order rational polynomial transformations, as with all other coordinate transformations, to make sure that the model used is not *over-parameterized*. Over-parameterization occurs when non zero coefficients are included in the transformation equations that

add distortion that does not exist in the "true" relationship between the coordinate systems.

By including and excluding different terms in the numerator and denominator, one can create an infinite number of variations of Eqs. (C-59). In addition, certain constraints on the equations can yield different transformations. For example, if all nonlinear terms and coefficients in the denominators are set to zero, and the following constraints are enforced: $a_1 = b_1$, $a_2 = -b_2$, the result is the 2D conformal coordinate transformation. The most common use of rational polynomials is in transforming three-dimensional object space coordinates to two-dimensional coordinates in satellite imagery. In this case, the right-hand sides of Eqs. (C-59) include three coordinates, usually functions of latitude longitude and height, and the left-hand sides are in terms of corresponding line and sample in the imagery.

C-11 Transformation Using Homogeneous Coordinates

In order to make transformation computations easier it is sometimes advantageous to embed points in a higher dimension by appending an extra coordinate. If $w \neq 0$, we say that $\square = (wx, wy, wz, w)$ are *homogeneous coordinates* for point $X = (x, y, z)$. Homogeneous coordinates are used extensively in computer graphics and computer vision. A benefit of using homogeneous coordinates is that if $w = 1$, we can apply a coordinate transformation using a single matrix. For example, rotation can be applied to a point $X = (x, y, z)^T$ with homogeneous coordinates $X = (x, y, z, 1)^T$ by

image

(C-60)

Translation can be applied by

image

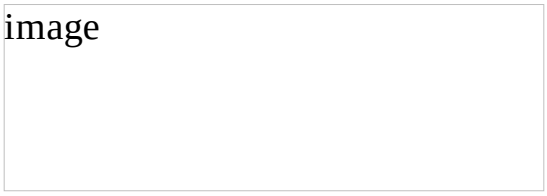
(C-61)

Finally, scale can be applied by

image

(C-62)

To apply a three-dimensional conformal coordinate transformation, the product of $\boldsymbol{T}\,\boldsymbol{S}$ and $\boldsymbol{R}$ can be used (with a single scale factor, $s = s_x = s_y = s_z$).



(C-63)

The following equation represents the same transformation as Eqs. (C-39) and (C-40) in homogeneous coordinates:

(C-64)

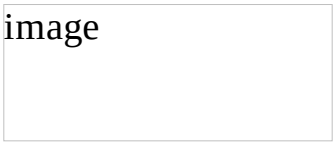
Also, note that

(C-65)

Similar to the 3D case, two-dimensional homogeneous coordinates for a point (x, y) are all points (wx, wy, w), where $w \neq 0$. Since two-dimensional homogeneous points are embedded into 3D space, we can easily interpret them graphically. Figure C-15 is a representation of the homogeneous coordinates of a two-dimensional point as a three-dimensional line. Notice that point (wx, wy, w) intersects the plane $z = w$ at (wx, wy).

FIGURE C-15 A three-dimensional line representing the set of homogeneous coordinates for a 2D point.

The 2D affine transformation of a point in homogeneous coordinates can be represented by a single matrix parameterized in the following way, using the same substitutions made in Eqs. (C-22):



(C-66)

Similarly, the 2D projective transformation can be represented using parameter definitions used in Eqs. (C-53):

(C-67)

Notice in Eq. (C-57) that the resulting third coordinate, *w*, may not equal 1. In this case, *X* and *Y* may be obtained by dividing the first two components by *w*. A result of this, the reverse transformation from (*X*, *Y*, 1) to (*x*, *y*, 1) may not be obtained by simply applying the inverse of the transformation matrix; one must take into account scaling by *w*. Also, note the distinction between the 2D projective transformation and the perspective projection transformation described in Sec. D-4.

Example C-5
An aerial photograph was taken over relatively flat ground. A 2D projective transformation was found from the ground to the image coordinate system with the following parameters:

a1	–0.002077558
b1	0.088303315
c1	–336.754
a2	–0.088445581
b2	–0.001989867
c2	360.534
a3	–0.000009327
b3	–0.000034412

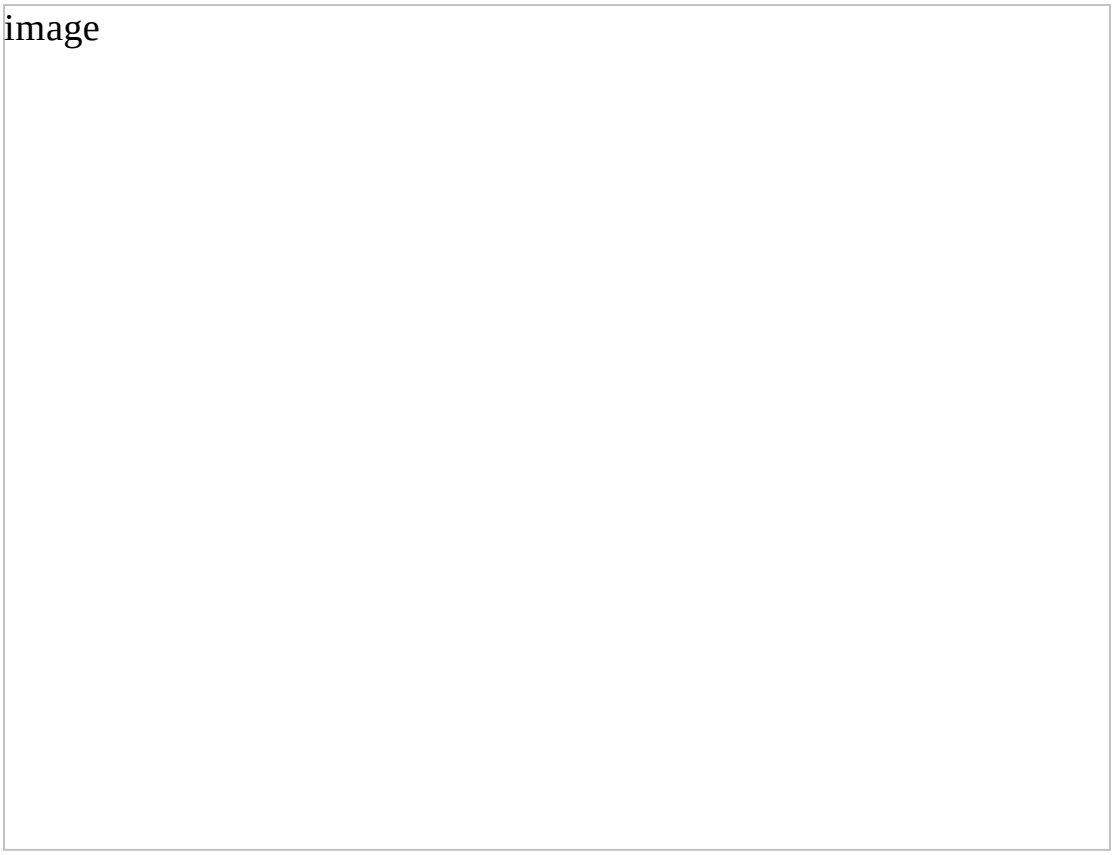
Using a spreadsheet program, parameterize the corresponding 2D projective rotation matrix, and transform, using homogeneous coordinates, the following points:

As a check, use a reverse transformation to transform the points back to the ground coordinate system.

Solution The transformation matrix in the form of Eqs. (C-57) can be multiplied by a matrix where the columns are comprised of the homogeneous coordinates of the points to be transformed yielding a matrix with columns containing the homogeneous coordinates of transformed points:

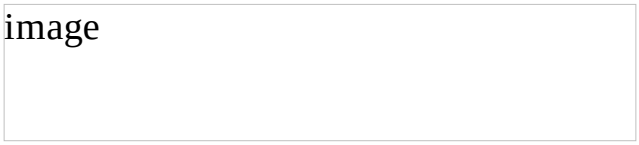
Dividing by the third component of each column yields the transformed coordinates of the points:

Similar to the method above, the inverse of the transformation matrix multiplied by the homogeneous coordinates of the transformed points yields the following:



Finally, divide each column vector by the third component, to yield the original matrix with

columns containing the original coordinates:



References

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Blais, J. A. R.: "Three-Dimensional Similarity," *Canadian Surveyor*, vol. 26, no. 1, 1972, p. 71.

Dewitt, B. A.: "Initial Approximations for the Three-Dimensional Conformal Coordinate Transformation," *Photogrammetric Engineering and Remote Sensing*, vol. 62, no. 1, 1996, p. 79.

Erio, G.: "Three-Dimensional Transformations for Independent Models," *Photogrammetric Engineering and Remote Sensing*, vol. 41, no. 9, 1975, p. 1117.

Fraser, C.S., G. Dial, J. Grodecki: "Sensor orientation via RPCs," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, 2006, p. 182.

Ghilani C.: *Adjustment Computations: Spatial Data Analysis*, Wiley & Sons, Hoboken, NJ, 2010.

Mikhail, E. M.: "Simultaneous Three-Dimensional Transformation of Higher Degree," *Photogrammetric Engineering*, vol. 30, no. 4, 1964, p. 588.

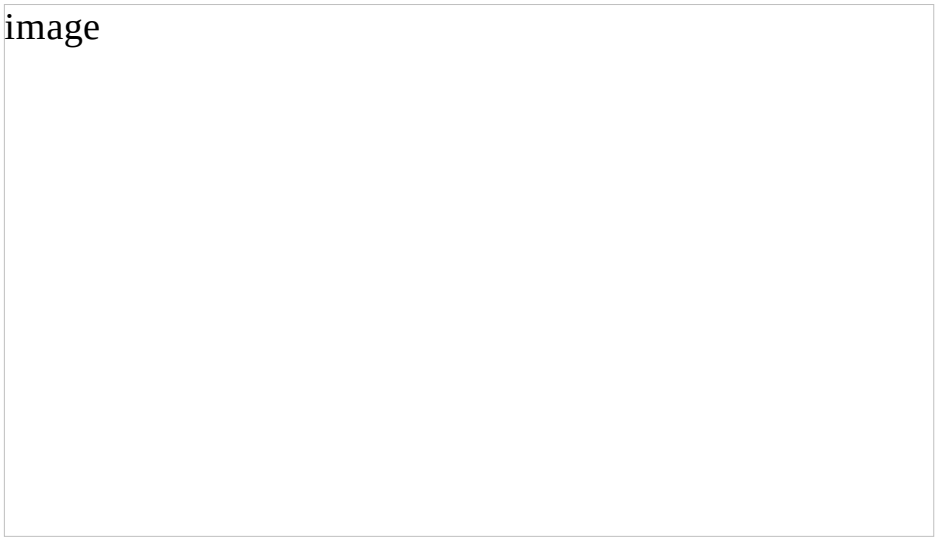
———: "Discussion Paper: Simultaneous Three-Dimensional Transformation," *Photogrammetric Engineering*, vol. 32, no. 2, 1966, p. 180.

Schut, G. H.: "Conformal Transformations and Polynomials," *Photogrammetric Engineering*, vol. 32, no. 5, 1966, p. 826.

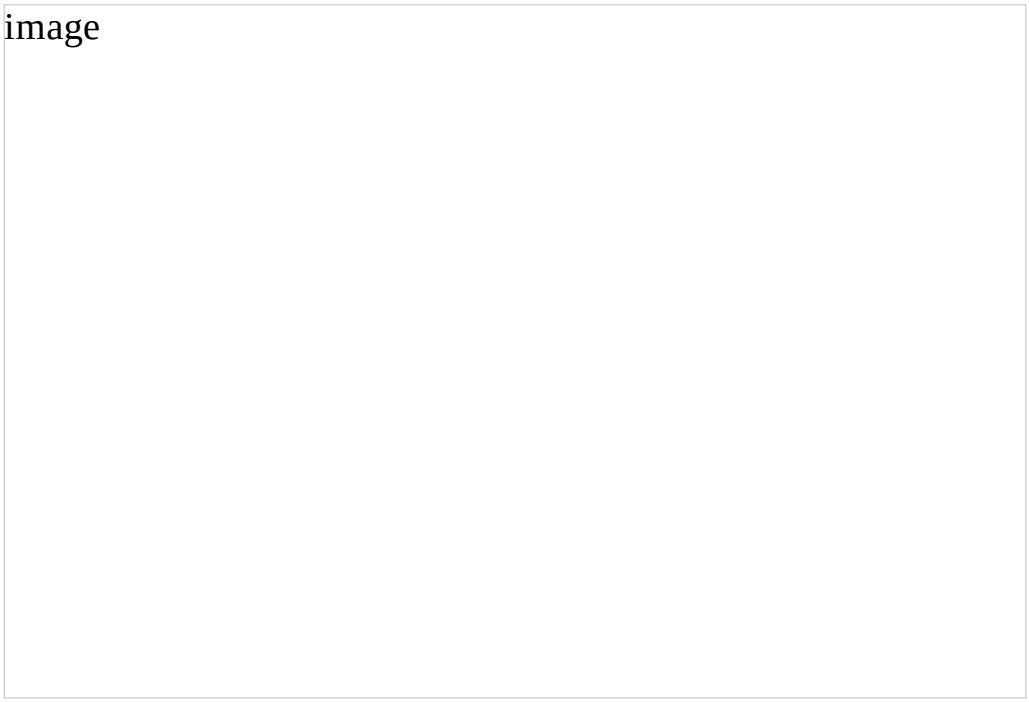
Yassa, G.: "Orthogonal Transformations," *Photogrammetric Engineering*, vol. 40, no. 8, 1974, p. 961.

Problems

C-1. The following table contains arbitrary x and y coordinates of a group of points that were measured from a digital map. The table also includes UTM coordinates for three of the points. Using a *two-dimensional conformal coordinate transformation*, and least squares, calculate UTM coordinates for the other points.



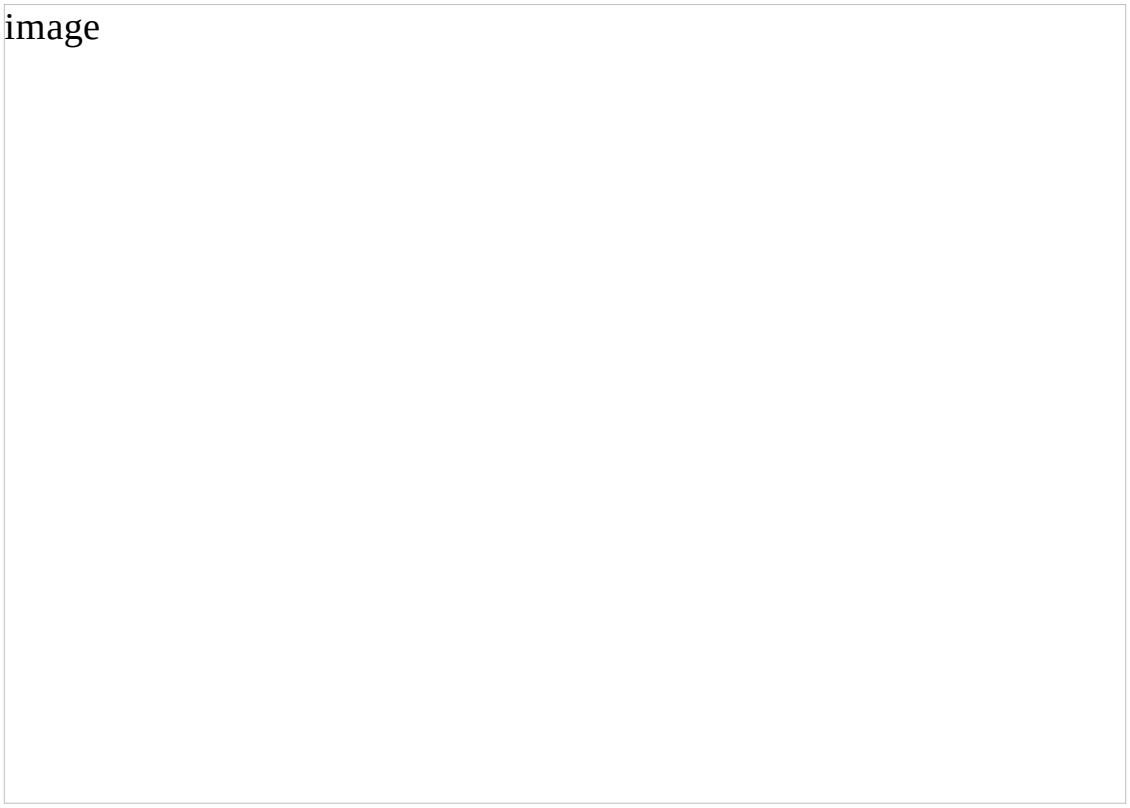
C-2. The following table contains *x* and *y* comparator coordinates for eight fiducial marks and three additional photo image points. The table also contains calibrated coordinates for the eight fiducials. Using a *two-dimensional affine transformation*, and least squares, calculate coordinates for the three image points in the fiducial axis system.



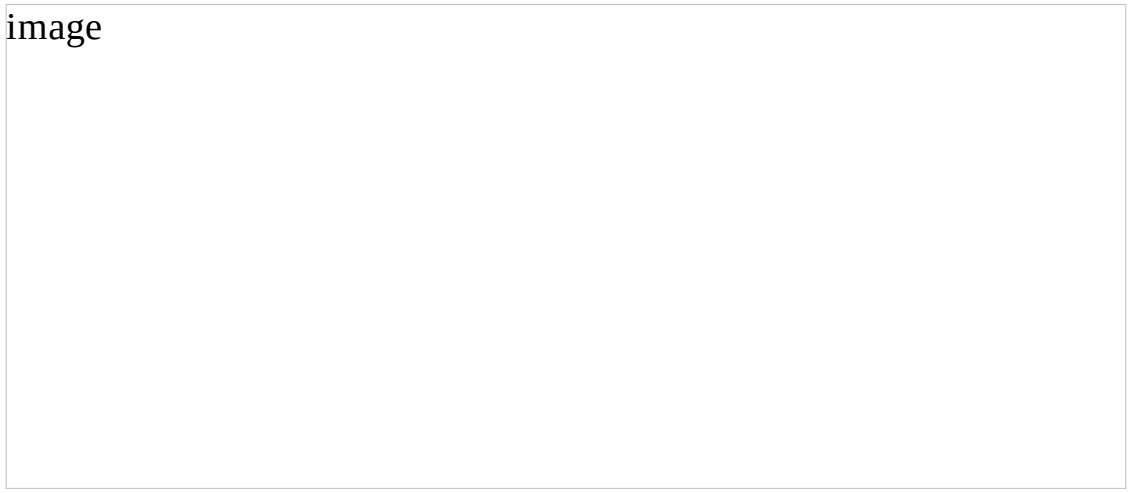
C-3. For the data of Prob. C-2, use the two-dimensional projective transformation and least squares to calculate coordinates for the three image points in the fiducial axis system.

C-4. Describe the concept of over-parameterization. What are the implications of this concept with respect to the results of Probs. C-2 and C-3?

C-5. Coordinates X_1, Y_1, and Z_1 for model I and X_2, Y_2, and Z_2 for model II of an independent model aerotriangulation are contained in the table below. Transform the model II coordinates into the model I coordinate system, using a three-dimensional conformal coordinate transformation and least squares (initial approximations of zero for omega and phi are sufficient).



C-6. Initial approximations are needed to solve for the 3D conformal coordinate transformation between two close-range stereomodels. Find initial approximations for all seven transformation parameters given the following control and arbitrary coordinates:



APPENDIX D

Development of Collinearity Condition Equations

D-1 Introduction

Collinearity, as illustrated in Fig. D-1, is the condition in which the exposure station of any photograph, an object point, and its photo image all lie on a straight line. The equations expressing this condition are called the *collinearity condition equations*. They are perhaps the most useful of all equations to the photogrammetrist.

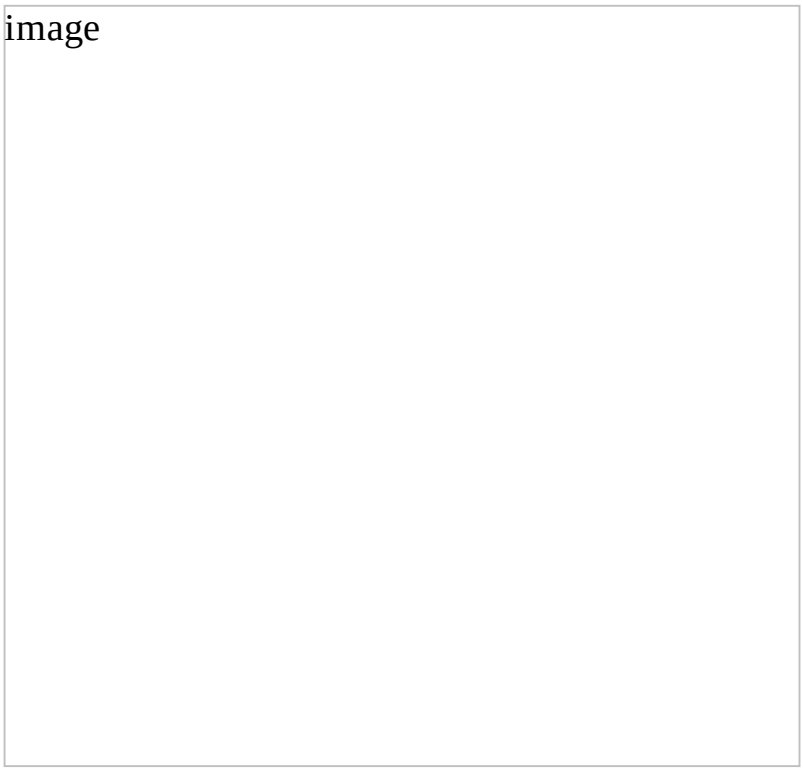


FIGURE D-1 The collinearity condition.

In Fig. D-2, exposure station L of an aerial photo has coordinates X_L, Y_L, and Z_L with respect to the object (ground) coordinate system XYZ. Image a of object point A, shown in a *rotated* image plane, has image space coordinates x'_a, y'_a, and z'_a, where the rotated image space coordinate system $x'y'z'$ is parallel to object space coordinate system XYZ. Initially, it is assumed that the principal point o is located at the origin of the xy photo coordinate system. A correction that compensates for this assumption is introduced at the end of the development.

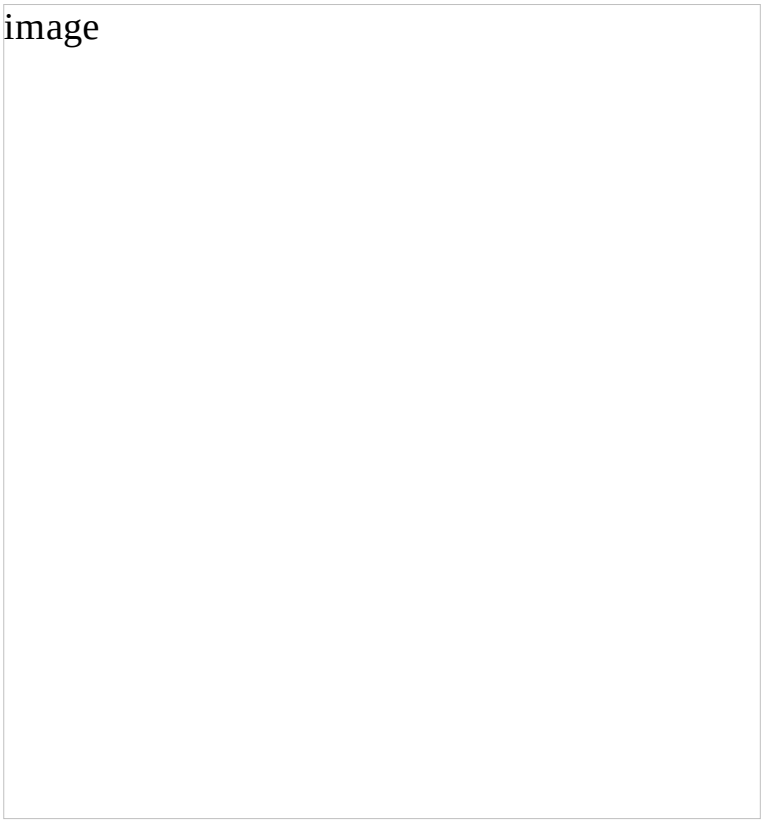


FIGURE D-2 Image coordinate system rotated so that it is parallel to the object space coordinate system.

D-2 Rotation in Terms of Omega, Phi, and Kappa

By means of the three-dimensional rotation formulas developed in App. C, an image point *a* having coordinates x_a, y_a, and z_a in a tilted photo such as that of Fig. D-1 may have its coordinates rotated into the $x'y'z'$ coordinate system (parallel to XYZ), as shown in Fig. D-3. The rotated image coordinates x'_a, y'_a, and z'_a, are related to the measured photo coordinates x_a and y_a, the camera focal length f, and the three rotation angles omega, phi, and kappa. The rotation formulas are Eqs. (C-32). For convenience they are repeated here:

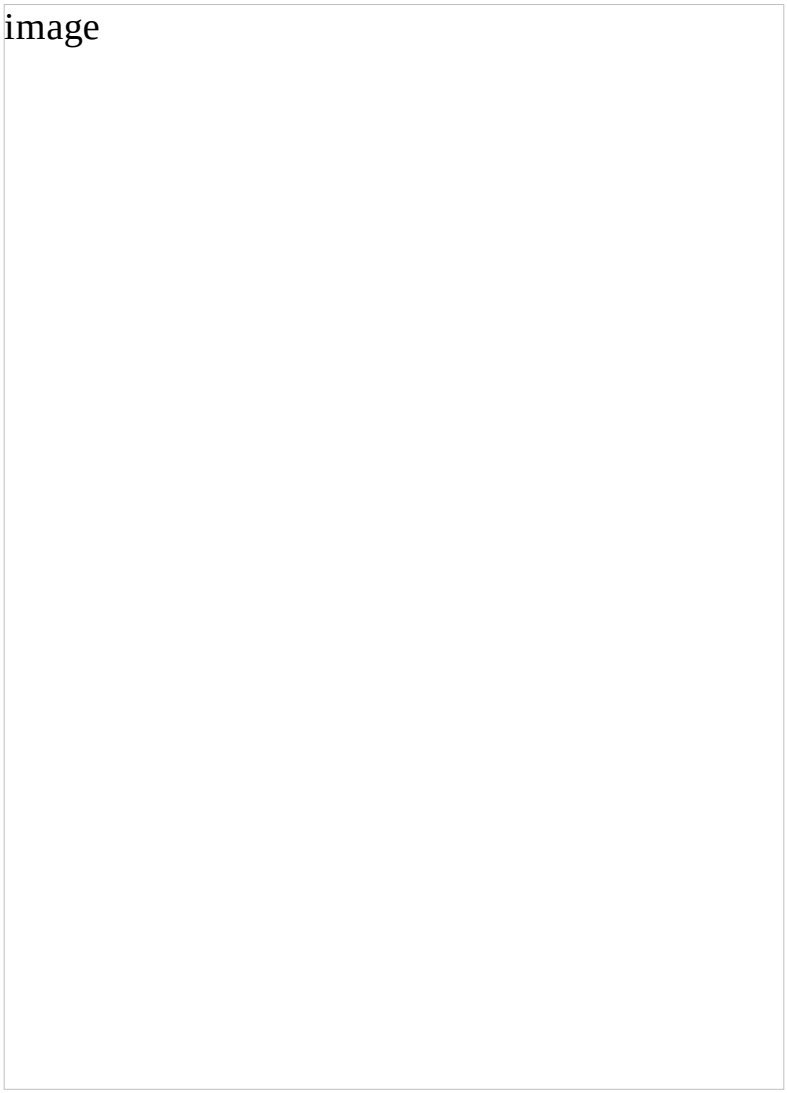
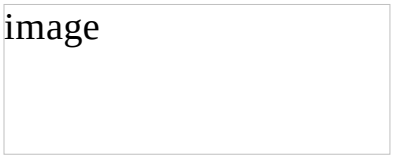


FIGURE D-3 Measurement *xyz* and rotated *x'y'z'* image coordinate systems.



(D-1)

In Eqs. (D-1) the *m's* are functions of the rotation angles omega, phi, and kappa. These functions are given in Eqs. (C-33). Also note in Fig. D-3 that the value for z_a is equal to $-f$.

D-3 Development of the Collinearity Condition Equations

The collinearity condition equations are developed from similar triangles of Fig. D-2 as follows:

Reducing gives

(*a*)

(*b*)

Also, by identity,

(*c*)

Substituting (*a*), (*b*), and (*c*) into Eqs. (D-1) gives

(D-2)

(D-3)

(D-4)

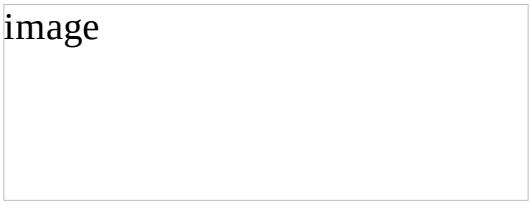
Factoring the term $z'_a/(Z_A - Z_L)$ from Eqs. (D-2) through (D-4), dividing (D-2) and (D-3) by (D-4), substituting $-f$ for z_a, and adding corrections for offset of the principal point (x_o, y_o), the following collinearity equations result:

(D-5)

(D-6)

D-4 Homogeneous Representation of the Collinearity Equations

As discussed in Sec. C-11, homogeneous coordinates can simplify the application of coordinate transformations. One can also use homogeneous representation to develop the collinearity equations by parameterizing a transformation matrix to project from three dimensions to two dimensions. Equation (D-7) contains the *perspective projection transformation matrix*, *P*:



(D-7)

Note that

(D-8)

Effectively, the perspective transformation matrix yields the coordinates of the intersection of a line from (X', Y', Z') to the origin with the plane at $z = 1/a$. Notice that the reverse transformation cannot be obtained because the matrix P is not invertible. This means that, although we can find a unique (x, y) for each point (X', Y', Z'), we cannot find a unique (X', Y', Z') given some (x, y) using only P.

If we consider (X'_A, Y'_A, Z'_A) the coordinates of a ground point in the camera coordinate system, we have the relationship shown in Eq. (D-9).

(D-9)

Since, the image coordinates (x, y) are where the line from (X'_A, Y'_A, Z'_A) to the origin (the perspective center in the camera coordinate system) intersects the plane at $z = –f$, we can choose the parameter a in the perspective transformation matrix as $–1/f$. Equation (D-10) is the homogeneous representation of the collinearity equations:

(D-10)

Note that

Multiplying both sides of Eq. (D-10) by this matrix and then dividing by their third components, yields Eqs. (D-5) and (D-6). Note that the scale difference, *w*, between the image coordinate system and the object coordinate system cancels out.

D-5 Linearization of the Collinearity Equations

Equations (D-5) and (D-6) are nonlinear and involve nine unknowns: the three rotation angles omega, phi, and kappa which are inherent in the *m*'*s*; the three exposure station coordinates X_L, Y_L, and Z_L; and the three object point coordinates X_A, Y_A, and Z_A. The photo coordinate measurements x_a and y_a are constant terms, as well as the calibration parameters x_o, y_o, and f which are considered to be constants in most applications of collinearity. The nonlinear collinearity equations are linearized by using Taylor's theorem. In linearizing them, Eqs. (D-5) and (D-6) are rewritten as follows:

(D-11)

(D-12)

where $q = m_{31}(X_A - X_L) + m_{32}(Y_A - Y_L) + m_{33}(Z_A - Z_L)$

$r = m_{11}(X_A - X_L) + m_{12}(Y_A - Y_L) + m_{13}(Z_A - Z_L)$

$s = m_{21}(X_A - X_L) + m_{22}(Y_A - Y_L) + m_{23}(Z_A - Z_L)$

According to Taylor's theorem, Eqs. (D-11) and (D-12) may be expressed in linearized form by taking partial derivatives with respect to the unknowns:

image

(D-13)

image

(D-14)

In Eqs. (D-13) and (D-14), F_0 and G_0 are functions F and G of Eqs. (D-11) and (D-12) evaluated at the initial approximations for the nine unknowns; the terms $(\partial F/\partial\omega)_0$, $(\partial G/\partial\omega)_0$, $(\partial F/\partial\varphi)_0$, $(\partial G/\partial\varphi)_0$, etc. are partial derivatives of functions F and G with respect to the indicated unknowns evaluated at the initial approximations; and $d\omega$, $d\varphi$, $d\kappa$, etc. are unknown corrections to be applied to the initial approximations. The units of $d\omega$, $d\varphi$, and $d\kappa$ are radians. Since the photo coordinates x_a and y_a are

measured values, if the equations are to be used in a least squares solution, residual terms must be included to make the equations consistent. The following simplified forms of the linearized collinearity equations include these residuals.

(D-15)

(D-16)

In Eqs. (D-15) and (D-16), J and K are equal to $x_a - F_0$ and $y_a - G_0$, respectively. The b's are coefficients equal to the partial derivatives. For convenience these coefficients are given below and on the next page. In these coefficients ΔX, ΔY, and ΔZ are equal to $X_A - X_L$, $Y_A - Y_L$, and $Z_A - Z_L$, respectively. Numerical values for these coefficient terms are obtained by using initial approximations for the unknowns.

D-6 Applications of Collinearity

The collinearity equations are applicable to the analytical solution of almost every photogrammetry problem. As examples, Sec. 11-6 describes their use in *space resection*, in which the six elements of exterior orientation of a tilted photograph are computed; and Sec. 11-10 explains how collinearity is applied in analytic *relative orientation*, which is necessary in analytically extending control photogrammetrically. Other applications are described elsewhere in this book. Regardless of the particular problem, an x equation [Eq. (D-15)] and y equation [Eq. (D-16)] are written for each point whose image or images appear in the photo or photos involved in the problem. The equations will contain unknowns, the number of which will vary with the particular problem. If the number of equations is equal to or greater than the number of unknowns, a solution is generally possible.

Initial approximations are needed for all unknowns, and these are usually easily obtained by making certain assumptions, such as vertical photography. The initial approximations do not have to be extremely close, but the closer they are to the unknowns, the faster a satisfactory solution will be reached; and the result is a savings in computer time.

In solving a system of collinearity equations of the form of Eq. (D-15) and (D-16) for any problem, the quantities that are determined are corrections to the initial approximations. After the first solution, the computed corrections are added to the initial approximations to obtain revised approximations. The solution is then repeated to find new corrections. This procedure is continued (iterated) until the magnitudes of the corrections become insignificant. A system of collinearity equations of the form of Eqs. (D-15) and (D-16) may be expressed in matrix form as

(D-17)

In Eq. (D-17), m is the number of equations; n is the number of unknowns; ${}_mV^1$ is the matrix of residual errors in the measured x and y photo coordinates; ${}_mA^n$ is the matrix of b's, the coefficients of the unknowns; ${}_nX^1$ is the matrix of unknown corrections to the initial approximations; and ${}_mL^1$ is the matrix of constant terms J and K. If the number of equations exceeds the number of unknowns, a least squares solution may be obtained for the most probable values for the unknowns by using matrix Eq. (B-12) or (B-13). Precision of the unknowns may be computed by applying matrix Eqs. (B-14) through (B-17).

D-7 Development of the Coplanarity Condition Equation

The coplanarity condition equation is developed by using simple vector calculus. Consider the three vectors in Fig. 11-2: the air base, B; the vector from the exposure station, L_1 the imaged point in photo 1, a_1 the object space coordinate system; and the vector from the exposure station, L_2 the imaged point in photo 2, a_2 the object space coordinates. The formulas for the last two vectors mentioned (from exposure stations to imaged points) are shown in Eqs. (D-18) and (D-19).

image

(D-18)

image

(D-19)

where M^T is the rotation matrix from the image coordinate system to the object coordinate system. If the rotation matrices and air base are correct for the two images, and image measurement noise is ignored, then the *triple-scalar product* of the three vectors is equal to zero. The triple-scalar product is the determinant of the matrix that has these three vectors as row entries. The equation for the 3 × 3-matrix determinant/triple-scalar product is shown in Eq. (D-20).

(D-20)

Due to incorrect exterior orientation parameters and random measurement errors, Eq. (D-20) is practically never equal to zero. In this case, the determinant equals the volume of a *parallelepiped* constructed using the three vectors. The parallelepiped is a three-dimensional shape where each edge has three other edges that are parallel and of equal length. Figure D-4 shows its construction.

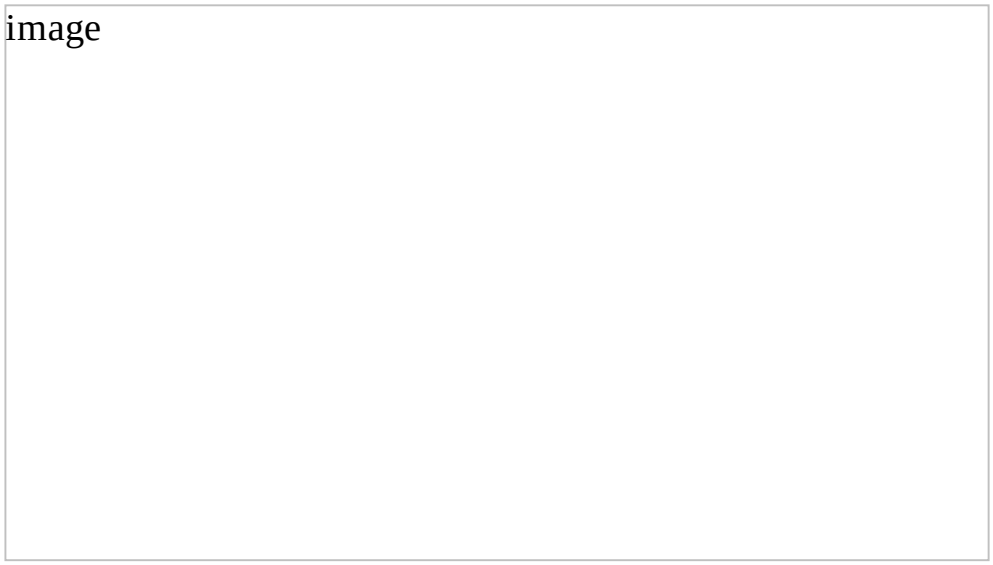


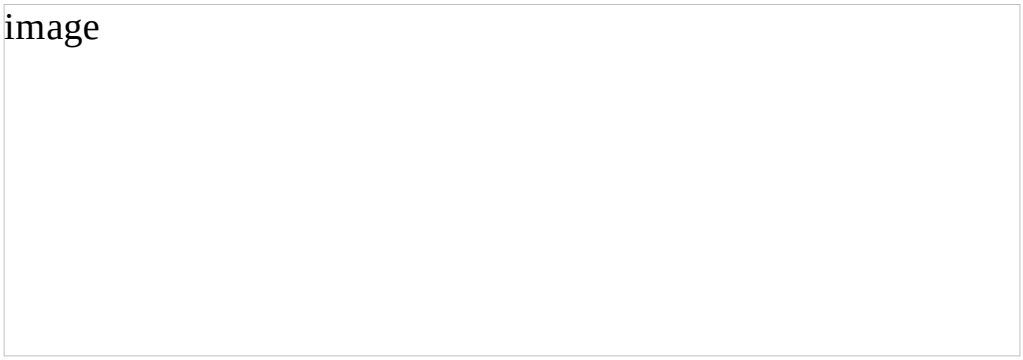
FIGURE D-4 Parallelepiped formed by three vectors used in the coplanarity condition equation.

By setting multiple observation equations Eq. (D-20) to zero, the goal when using coplanarity, in a least squares sense, is to find the exterior orientation parameters that minimize the sum of the squares of the volumes of the resulting parallelepipeds. Note that no object space coordinates of imaged points are used in the coplanarity equation. This means that the object space control must come from some of the exterior orientation parameters. A common application of coplanarity is to fix seven of the twelve exterior orientation parameters of two images to perform relative orientation as described in Sec. 11-5. However, observations of all exterior orientation parameters can be included and adjusted based on the coplanarity equations of one or more points. Similar to collinearity, the coplanarity condition equation is nonlinear and must be linearized using Taylor's theorem.

Since using coplanarity involves calculating determinants, oftentimes very large numbers are involved. This can lead to ill-conditioning in implementation. This is one reason that collinearity is used more often than coplanarity. However, there are various methods to handle ill-conditioning, such as the one described in Example B-6.

D-8 Linearization of the Coplanarity Equation

Equation (D-20) has twelve unknowns, two sets of six exterior orientation parameters associated with the images involved. The linearized form of Eq. (D-20) is as follows:



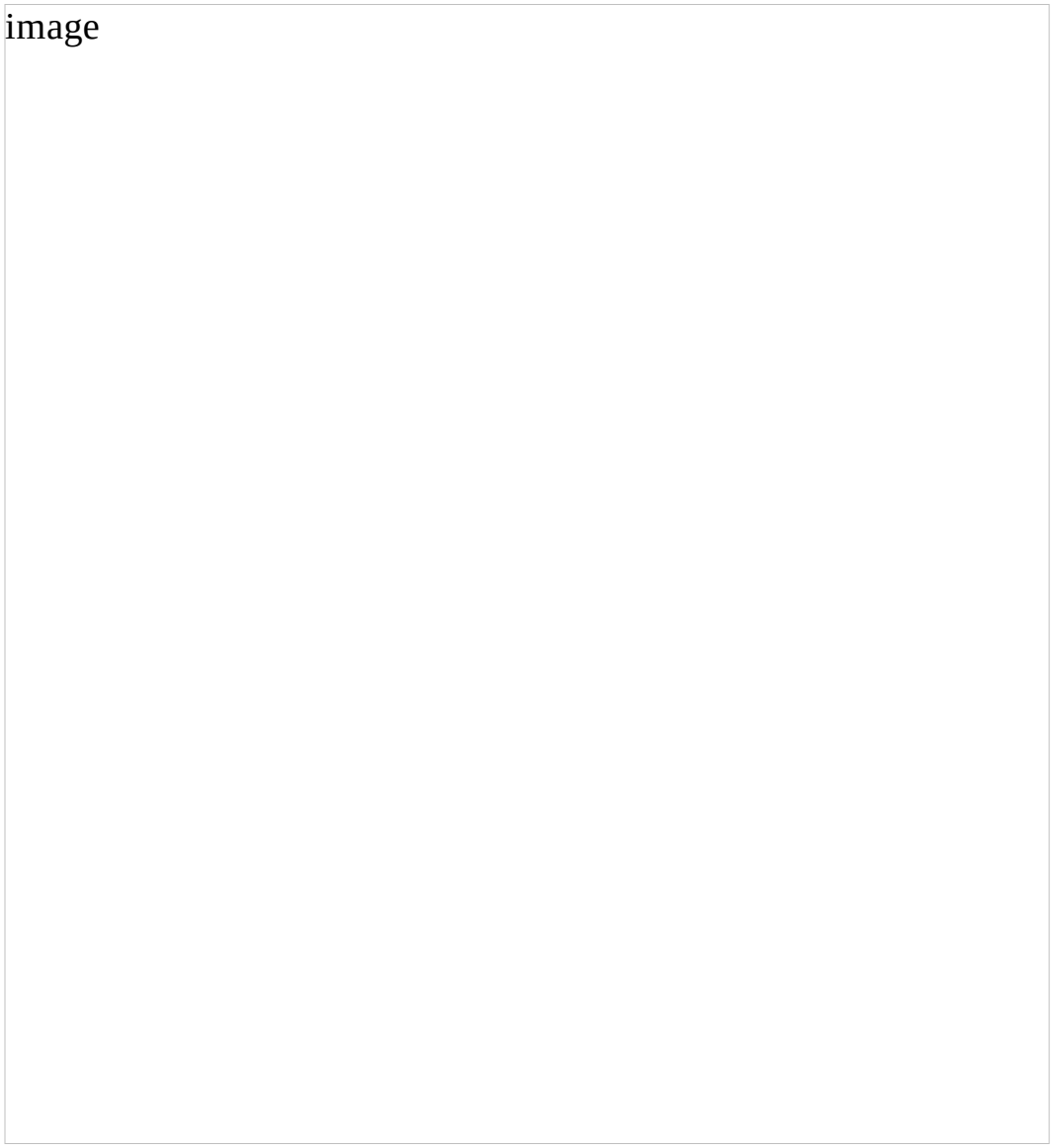
(D-21)

The term H_0 is Eq. (D-20) evaluated at initial approximations of the twelve unknowns. The terms , etc. are the partial derivatives of Eq. (D-20) with respect to the unknown exterior orientation parameters evaluated at the initial approximations. Note that when using coplanarity to perform relative orientation of two images, only five of the partial derivatives in Eq. (D-21) will have nonzero values, since seven parameters are considered constants. The terms , etc. are corrections to be applied to the initial approximations. As with the collinearity equations, a residual term must be included for consistency. However, since the coplanarity equation includes multiple measurements, x_1, y_1, x_2, and y_2, residuals are in terms of the entire equation. The simplified form of the linearized coplanarity equation is shown in Eq. (D-22).

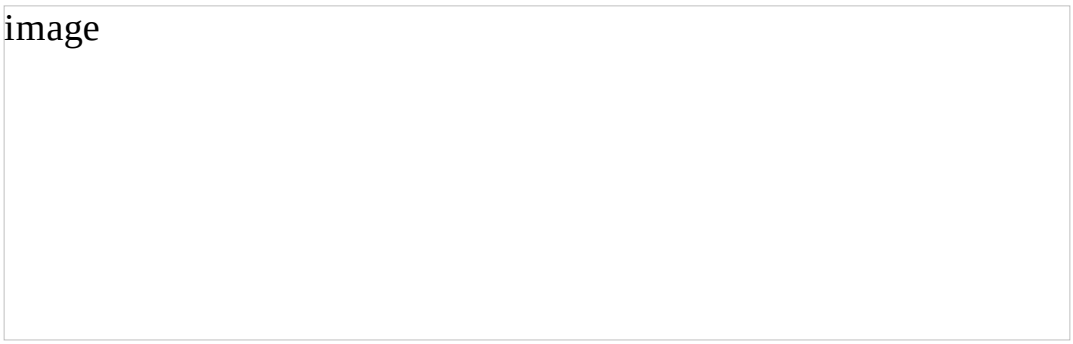
(D-22)

The partial derivative coefficients for Eq. (D-22) are given as follows:

image



where



D-9 Rotation in Terms of Azimuth, Tilt, and Swing

Instead of using omega, phi, and kappa as the rotation angles for transforming the tilted photo coordinates into an $x'y'z'$ coordinate system parallel to the ground system, the rotation angles of azimuth, tilt, and swing may be used. A tilted photo illustrating the azimuth α, tilt t, and swing s angles is shown in Fig. D-5. In the figure the photo principal plane intersects the datum plane along line N_dP_d. The rotation formulas are developed by initially assuming an $x'y'z'$ coordinate system parallel to XYZ and then, by means of rotations, converting it to the xyz photo measurement system. The origins of both image coordinate systems are taken at exposure station L.

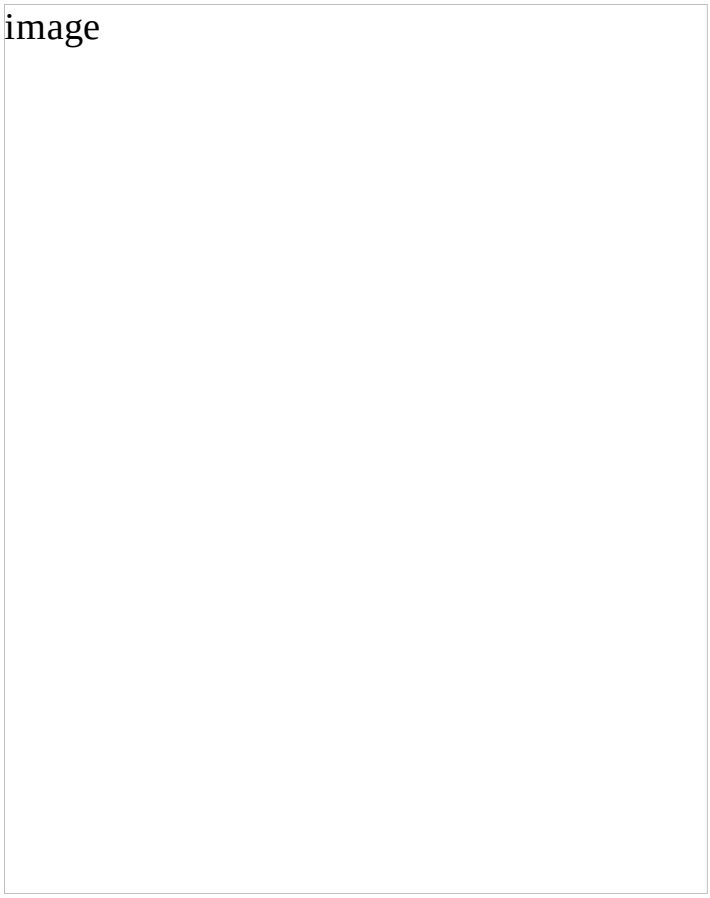


FIGURE D-5 Azimuth, tilt, and swing rotation angles.

The rotation equations are developed in a sequence of three separate two-dimensional rotations. The $x'y'z'$ coordinate system is first rotated about the z' axis through a clockwise angle α to create an $x^{\alpha}y^{\alpha}z^{\alpha}$ coordinate system. After rotation the y^{α} axis will be in the principal plane of the photo. With reference to Fig. D-6*a*, the coordinates of any point in the $x^{\alpha}y^{\alpha}z^{\alpha}$ system are

FIGURE D-6 Rotation in azimuth, tilt, and swing. (*a*) First rotation. (*b*) Second rotation. (*c*) Third rotation.

(D-23)

The second rotation is a counterclockwise rotation t about the x^{α} axis to create an $x^{\alpha t}y^{\alpha t}z^{\alpha t}$ coordinate system. After rotation, the $x^{\alpha t}$ and $y^{\alpha t}$ axes are in the plane of the tilted photograph. With reference to Fig. D-6*b*, the coordinates of any point in the $x^{\alpha t}y^{\alpha t}z^{\alpha t}$ system are

(D-24)

The third rotation is about the $z^{\alpha t}$ axis through the counterclockwise angle θ. Angle θ is defined as

This third rotation creates an $x^{\alpha t\theta}y^{\alpha t\theta}z^{\alpha t\theta}$ coordinate system which coincides with the *xyz* tilted photo system. With reference to Fig. D-6*c*, the coordinates of any point in the *xyz* system are

(D-25)

Because sin θ equals –sin *s*, and cos θ equals –cos *s*, these substitutions may be made into Eqs. (D-25), from which

(D-26)

Substituting Eqs. (D-23) into Eqs. (D-24), in turn substituting into Eqs. (D-26), and factoring, the following expressions for the *x*, *y*, and z coordinates of any point are obtained:

(D-27)

In Eqs. (D-27), the *m*'*s* are

(D-28)

D-10 Collinearity Equations Using Azimuth-Tilt-Swing Rotation

By simply substituting Eqs. (D-28) for the *m's* into Eqs. (D-5) and (D-6), collinearity equations are obtained which include azimuth, tilt, and swing as unknowns instead of omega, phi, and kappa. By applying Taylor's theorem, these azimuth, tilt, and swing equations can be linearized and used to solve photogrammetry problems analytically. More frequently, however, the omega-phi-kappa equations are used, and if the azimuth, tilt, and swing angles are desired, they are determined from omega, phi, and kappa as described in the next section.

D-11 Converting from One Rotation System to the Other

Although the azimuth-tilt-swing expressions [Eqs. (D-28)] for the *m's* differ from their corresponding omega-phi-kappa expressions [Eqs. (C-33)], for any given tilted photo their numerical values are equal. This is true because the *m's* are actually direction cosines relating image and object coordinate systems as described in Sec. C-7. Because of their equality, corresponding *m's* may be set equal to each other; for example, $m_{11} = \cos\varphi \cos\kappa = -\cos\alpha \cos s - \sin\alpha \cos t \sin s$. This equality of *m's* enables conversions to be made between the omega-phi-kappa and azimuth-tilt-swing systems.

Before discussing the conversion method, it is necessary to establish limits on the ranges of values for the two sets of angles. This is required since if each angle is allowed to have a full 360° range, any particular rotation configuration will have two equally valid sets of angles in either the azimuth-tilt-swing or the omega-phi-kappa system. For example, azimuth = 0°, tilt = 5°, and swing = 0° will generate the same rotation matrix as azimuth = 180°, tilt = –5°, and swing = 180°. Ranges for azimuth, tilt, and swing which avoid this multiple-definition problem are

and ranges for omega, phi, and kappa are

Although not necessary, any of the ranges for azimuth, swing, omega, or kappa could be chosen as 0° to 360°, if desired.

If omega, phi, and kappa for a particular photograph are known, numerical values for the *m*'*s* can be calculated by Eqs. (C-33) and the tilt, swing, and azimuth determined from the following:

(D-29)

(D-30)

(D-31)

In Eqs. (D-30) and (D-31) it is essential that a full-circle inverse tangent function (such as atan2) be used so that the full ranges for s and α can be determined. In the rare case where tilt is exactly 0° or 180°, both the numerator and denominator in each of Eqs. (D-30) and (D-31) will equal zero, resulting in invalid results from a full-circle inverse tangent function. (Note that if the denominator is zero but the numerator is nonzero, a properly implemented full-circle inverse tangent function will return valid results.) In this situation where tilt is exactly zero, no principal line exists, and swing and azimuth are undefined. However, by arbitrarily defining azimuth to be equal to zero, a value for swing can be obtained from the rotation matrix. The original definitions of swing and azimuth will no longer apply (see Sec. 10-2); however, the resulting values can still be properly used in photogrammetric equations. When the tilt is exactly 0° (or 180°), the values for swing and azimuth can be obtained by

(D-32)

(D-33)

If the azimuth, tilt, and swing are known for a particular photo, conversion to omega, phi, and kappa is also readily made as follows:

(D-34)

(D-35)

(D-36)

Once again, it is essential that a full-circle inverse tangent function be used with Eqs. (D-35) and (D-36) so that values of omega and kappa in the proper ranges are computed. In the rare case where phi is exactly ±90°, both the numerator and denominator in each of Eqs. (D-35) and (D-36) will be

zero, and the values for omega and kappa will be undefined. By giving an arbitrary definition of zero for omega as indicated by Eq. (D-37), Eq. (D-38) may then be used to compute kappa.

(D-37)

(D-38)

Reference

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004.

Problems

D-1. State the condition of collinearity in photogrammetry.

D-2. Explain why linearized collinearity equations must be iterated a number of times before a satisfactory solution is achieved.

D-3. Explain why the parallelepiped described in Sec. D-7 has zero volume when the exterior orientation parameters of the corresponding two images are correct (or correct relative to each other).

D-4. Given the following values, compute the photo coordinates x_a and y_a, using Eqs. (D-5) and (D-6). Express your answers to the nearest 0.001 mm.

D-5. Using matrix multiplication, show in steps (each consecutive matrix product) that Eq. (D-10) is equivalent to Eqs. (D-5) and (D-6). Be sure to substitute the vector defined in Eq. (D-9).

D-6. Convert the following values of azimuth, tilt, and swing to omega, phi, and kappa. Express your answers in decimal degrees to the nearest thousandth of a degree.

(*a*) $\alpha = 30.000°$, $t = 0.500°$, $s = -165.000°$
(*b*) $\alpha = -145.000°$, $t = 97.000°$, $s = 23.000°$
(*c*) $\alpha = 90.000°$, $t = 90.000°$, $s = -85.000°$

D-7. Convert the following values of omega, phi, and kappa to azimuth, tilt, and swing. Express your answers in decimal degrees to the nearest thousandth of a degree.

(*a*) $\omega = 7.800°$, $\phi = -1.100°$, $\kappa = 0.000°$

(*b*) $\omega = -115.000°$, $\phi = 37.000°$, $\kappa = -10.000°$
(*c*) $\omega = 0.000°$, $\phi = 0.000°$, $\kappa = 0.000°$

APPENDIX E
Digital Resampling

E-1 Introduction

The acquisition of a digital image involves discrete sampling of a continuous analog signal, for example, the reflected energy from the ground. These digital samples are made in a (distorted) grid pattern, with each grid cell or pixel containing a digital number (DN) representing the lightness or darkness at its corresponding ground location. When a digital image is acquired, no attempt is made to have the pixels line up with any particular map projection coordinates. It is therefore necessary to perform resampling to obtain a digital sample at an intermediate (i.e., fractional) row, column location. Resampling involves interpolation between existing pixels (DNs) to synthesize pixels that correspond to fractional locations as illustrated in Fig. E-1. Determination of the appropriate fractional locations is often the result of a coordinate transformation (see App. C).

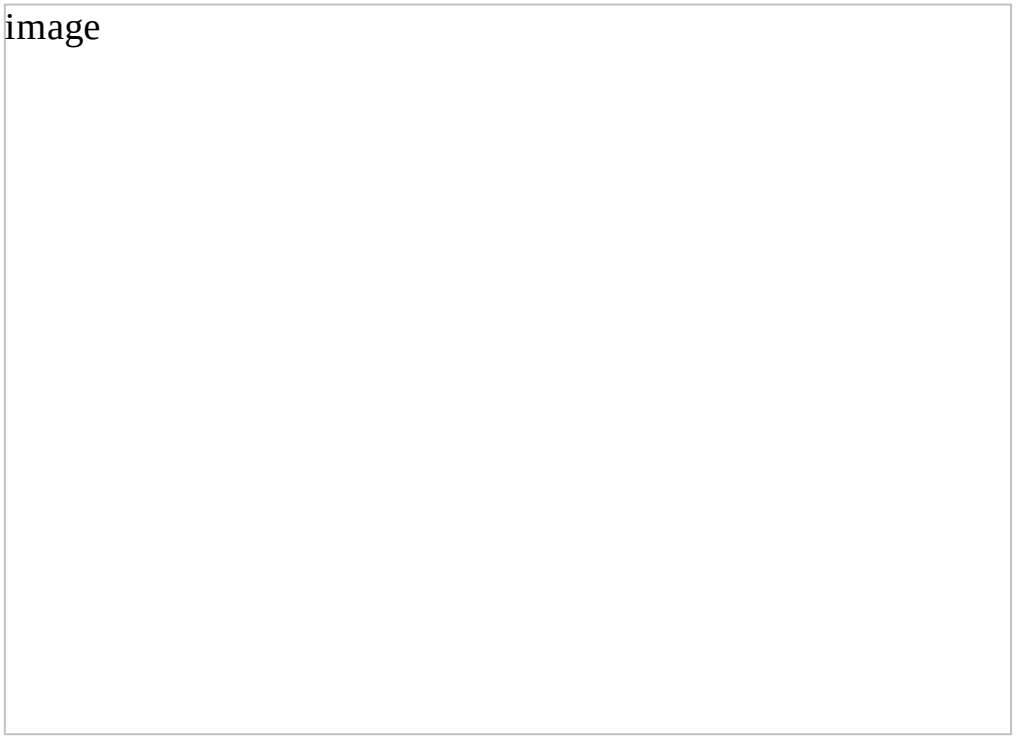


FIGURE E-1 Relationship between pixels from the originally sampled image and the resampled image.

There are several techniques available for resampling digital images, although three particular ones are by far, most prevalent. They are known as *nearest-neighbor* interpolation, *bilinear* interpolation, and *bicubic* interpolation. Other, more computationally intensive techniques are generally not employed since they tend to be sensitive to sensor noise which exists in digital imagery.

E-2 Nearest Neighbor

The nearest-neighbor interpolation is simplest of the three. As its name implies, the DN chosen will be that of the image pixel whose center is closest to the center of the grid cell. From a computational standpoint, all that is required is to round off the fractional row and column values to the nearest integral value. Figure E-2 shows the DNs for a 4 × 4 subarea from a digital image. A pixel is superimposed at a fractional location (R = 619.71, C = 493.39). Rounding these values to the nearest integer yields 620 and 493 for the row and column indices, respectively. Thus, the resampled value is 56.

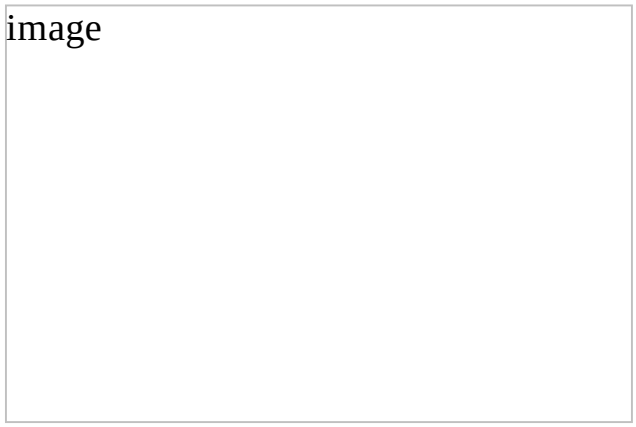


FIGURE E-2 A 4 × 4 subarea of image pixels with superimposed grid cell at a fractional location.

E-3 Bilinear Interpolation

A second resampling method, which involves greater computational complexity, is bilinear interpolation. In this approach, the four surrounding pixels are selected, and linear interpolation is performed in two directions. This is illustrated by example, using the values from Fig. E-2. First, linearly interpolated values DN_1 and DN_2 are computed along rows 619 and 620, respectively. Equations (E-1) and (E-2) illustrate the calculation.

(E-1)

(E-2)

Next, a linear interpolation is performed in the column direction, yielding the result (DN) given in Eq. (E-3).

(E-3)

Finally, since DNs are generally integers, the value from Eq. (E-3) is rounded to 59. Note that this differs from the DN value of 56 obtained by using the nearest-neighbor method.

E-4 Bicubic Interpolation

Bicubic interpolation, also known as *cubic convolution*, is a third resampling technique which is commonly used. Explanation of this technique requires a little background in sampling theory. First,

an assumption is made that the original signal has been sampled above the *Nyquist rate*, which is generally satisfied for imaging sensors. The Nyquist rate is, in essence, the sampling frequency required to faithfully record the highest (spatial) frequency content of the scene. Given this assumption, the "sinc" function allows an (almost) exact reconstruction of the original scene. The form of the sinc function is shown in Fig. E-3. (Note that in this figure, the argument for the sine function is in radians.) If the images had an infinite number of rows and columns, and all pixels were used for the interpolation, the sinc function would yield a perfect reconstruction. Practicality, however, dictates that interpolations be carried out using only small neighborhoods surrounding the interpolated pixel. A cubic spline approximation to the sinc function is the form generally used for bicubic interpolation. The shape of the spline is given in Fig. E-4 while Eqs. (E-4) through (E-6) express the functional relationship. For comparison, Fig. E-4 also shows the shape of nearest-neighbor and bilinear interpolations expressed in the form of similar weighting functions. Note that the cubic spline most nearly approximates the sinc function of Fig. E-3, whereas bilinear and nearest-neighbor interpolations are less consistent approximations.

FIGURE E-3 Form of the sinc function.

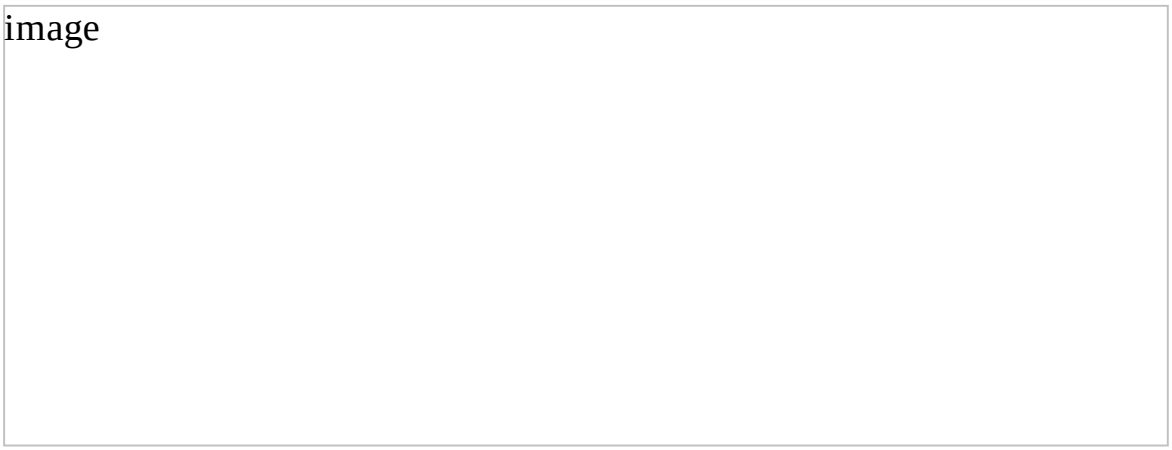


FIGURE E-4 Shape of the bicubic, bilinear, and nearest-neighbor interpolations, in the form of weighting functions $h(x)$.

(E-4)

(E-5)

(E-6)

where a = free parameter equal to slope of weighting function at $x = 1$ (generally $a = -0.5$ yields good results)

x = absolute value of difference between whole-number row or column and fractional row or column

The computational process is analogous to that of bilinear interpolation in that it is performed first along the rows and then down the single, fractional column. (The procedure is demonstrated with a numerical example below.) The computations are conveniently expressed in matrix form, as shown in Eq. (E-7). In this equation, the R and C matrices consist of coefficients derived from Eqs. (E-4) and (E-5), and the D matrix contains the digital numbers from the 4×4 neighborhood surrounding the interpolated pixel.

image

(E-7)

Interpolating across the rows (based on the fractional column position) is done by forming the product DC. Subsequently, R is multiplied by the product to obtain the final interpolated value. In fact, the sequence of multiplications does not matter; they can be performed from left to right in Eq. (E-7).

Referring to the example of Fig. E-1, the bicubic interpolation computation procedure begins by computing the elements of the R matrix. Given the fractional row location (619.71), an interpolation weighting function will be computed for the two rows above (618 and 619) and the two rows below (620 and 621). The distance x from the fractional row is determined for each of the four surrounding rows, and the corresponding function [Eq. (E-4) or (E-5)] is selected. The result of this computation is listed in Table E-1. In a similar fashion, the elements of matrix C are computed, as shown in Table E-2.

image

TABLE E-1 Row Interpolation Weight Matrix R Computed for Example of Fig. E-1. Slope Value of $a = -0.5$ was Used

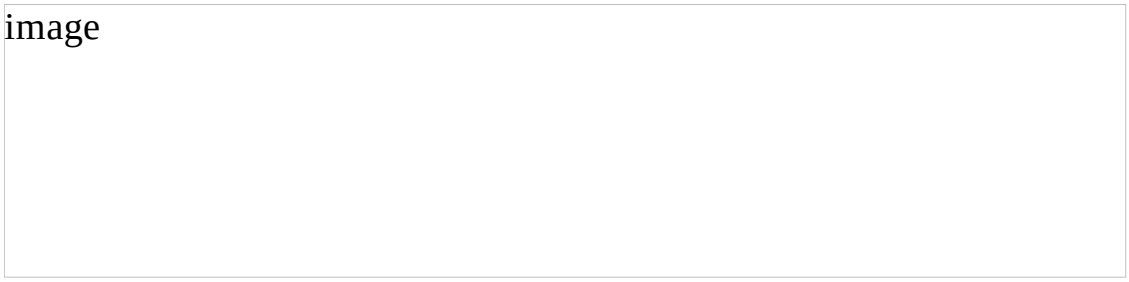


TABLE E-2 Column Interpolation Weight Matrix *C* Computed for Example of Fig. E-1. Slope Value of $a = -0.5$ was Used

Following these preliminary calculations, the matrix product from Eq. (E-7) is formed. This product, using the example values, is shown in Eq. (E-8). The resultant value of 60.66 is then rounded to 61, the nearest integer. Note that again this DN value differs from both the 56 of nearest-neighbor interpolation and the 59 of bilinear interpolation.

(E-8)

E-5 Discussion of the Three Methods

The primary advantage of the nearest-neighbor method is that it is the fastest of the three techniques in terms of computational time. It also has the advantage of not modifying the original image data, which is important if remote-sensing image classification will be performed. However, since a continuous interpolation is not being performed, the resultant appearance can be somewhat jagged or blocky.

The primary advantage of bilinear interpolation is the smoother appearance of the result. This appearance is slightly compromised by the fact that some high-frequency detail is filtered out. In other words, edges in the scene are slightly less distinct. In terms of computational time, bilinear interpolation is slower than nearest-neighbor interpolation but faster than bicubic interpolation.

Bicubic interpolation is the most rigorous resampling technique of the three on the basis of signal processing theory. It achieves the smooth appearance without sacrificing as much high-frequency (edge) detail as is characteristic of the bilinear interpolation. The reason for this is illustrated in Fig. E-5. This figure shows a profile of four consecutive samples along a single row of a digital image. The curve represents a smooth (cubic spline) interpolation of image brightness based on the samples at the four dots. The dashed line represents the linear interpolation between the two center pixels. For the indicated fractional column, the bicubic interpolation is able to capture the trend of the signal better than bilinear interpolation. By capturing trends in the signal, the bicubic interpolation retains high frequency information that would be clipped by the bilinear interpolation method. However, this enhanced appearance comes at a penalty in terms of computational time.

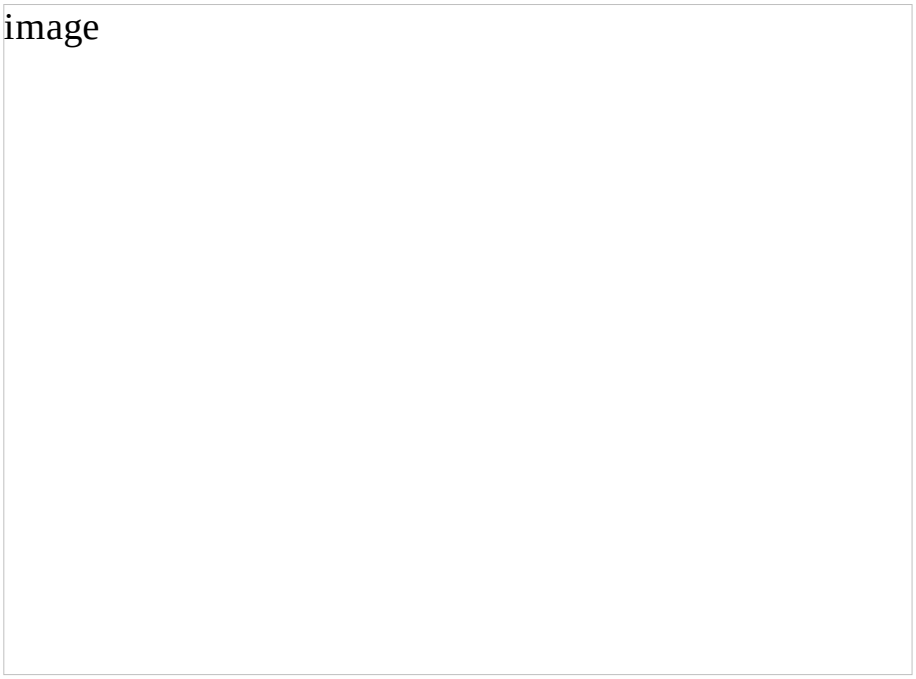


FIGURE E-5 Bicubic interpolation captures signal trend more closely than bilinear interpolation.

References

American Society for Photogrammetry and Remote Sensing: *Manual of Remote Sensing*, 3d ed., vol.1, Bethesda, MD, 1998.

Billingsley, F. C.: "Review of Image Processing Fundamentals," *Digital Image Processing, SPIE*, vol. 528, 1985.

Moik, J. G.: *Digital Processing of Remotely Sensed Images*, NASA SP-431, Government Printing Office, Washington, 1980.

Problems

E-1. Using the data from the example presented in this appendix, calculate a bicubic interpolation using a slope parameter $a = -1$.

E-2. Using the 4 × 4 array of digital numbers from Fig. E-1, compute a resampled pixel value by each of the three given methods corresponding to row = 619.35, column = 493.78. Use the value of $a = -0.5$ for the slope parameter of the bicubic interpolation method.

APPENDIX F

Conversions Between Object Space Coordinate Systems

F-1 Introduction

In Chap. 5, descriptions of four fundamental object space coordinate systems were given. In this appendix, mathematical formulas which can be used to convert coordinates between the various systems are presented. *Conversions* differ from *transformations* in that conversions are exact mathematical processes which simply change the mathematical representation of positions from one form to another. Changing two-dimensional rectangular coordinates to polar coordinates is an example of a conversion. A transformation, on the other hand, is not an exact mathematical process per se, but is based on measurements and therefore contains errors. This is the key distinguishing characteristic between conversions and transformations. If a different set of measurements is used, a transformation may yield different results; but since a conversion is not based on measurements, it will always be consistent.

Figure F-1 illustrates the sequence in which the conversions are accomplished. At the top of this figure are map projections which assign plane coordinates XY based on developable surfaces. These coordinates can be converted to geodetic coordinates of latitude ϕ and longitude λ. The two-way arrow between the two coordinate systems in the figure indicates that the conversion can also be made in reverse, i.e., conversion from geodetic to map projection coordinates. Following downward in Fig. F-1, the next conversion represented is from geodetic latitude ϕ, longitude λ, and ellipsoid height h to geocentric coordinates X, Y, and Z. Again, this conversion can be made in reverse. At the bottom of the figure, the conversion of geocentric coordinates to local vertical coordinates X_l, Y_l, and Z_l is illustrated; and as before, the conversion can be reversed. The conversions will always follow the path indicated by Fig. F-1. If conversion is to be made from map projection coordinates to local vertical, for instance, calculation of geodetic and geocentric coordinates will be performed as intermediate steps.

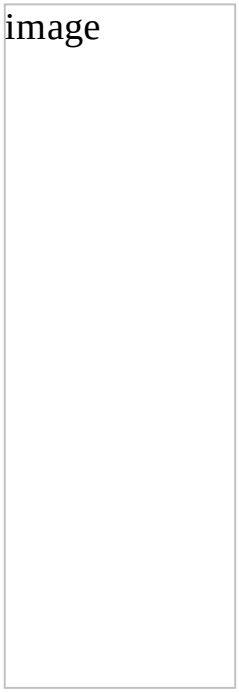


FIGURE F-1 Object space coordinate conversion sequence.

In many cases, values of coordinates in the aforementioned systems may have as many as 10 or 11 significant figures. When performing coordinate conversions using values this precise, it is essential to carry a large number of significant figures in the computations. A hand calculator having 10 significant figures may give marginal results. It is suggested that solutions be programmed on a computer using double-precision variables or solved with a spreadsheet having at least 12 significant figures. Off-the-shelf computer programs that perform these coordinate conversions generally use suitable precision for internal calculations, but this should be verified to ensure accurate results. *Note:* The examples in this appendix were solved using 16 significant figures.

F-2 Standard Ellipsoid Constants for GRS80

To define the size and shape of the ellipsoid, at least two constants are required. Generally, the semimajor axis of the ellipsoid a and the flattening f are used. From these two defining constants, other ellipsoid parameters can be derived.

The examples in this appendix are based on the GRS80 ellipsoid. Since many of the conversions presented in this chapter require values for standard ellipsoid constants, they are presented here. From Table 5-1 the values for a and $1/f$ are

image

From these quantities, the values of b, e, e^2, and e'^2 can be computed as follows:

From Eq. (5-2):

From Eq. (5-4):

image

From Eq. (5-6):

F-3 Conversion Between Geodetic and Geocentric Coordinates

Conversion between geodetic $\phi\lambda h$ and geocentric XYZ is a three-dimensional process. Figure 5-5 illustrates the relationship between the two systems. Since geodetic coordinates are related to the

reference surface used, certain ellipsoid constants (a, e^2, and e'^2) are required in order to achieve the conversion. To convert a position expressed in geodetic coordinates to geocentric coordinates, the following equations, which can be found in standard geodesy texts, are used:

(F-1)

(F-2)

(F-3)

In Eqs. (F-1) through (F-3), the value N is the length of the normal to the ellipsoid at latitude ϕ and is computed by

(F-4)

Conversion from geocentric coordinates to geodetic coordinates may be accomplished by the following approach. First, longitude is calculated by dividing Eq. (F-2) by (F-1) and taking the inverse tangent as follows:

(F-5)

In Eq. (F-5) it is essential that the full-circle inverse tangent be used since longitude can range from -180° to +180°.

Latitude and height are not as easily isolated for direct solution. A number of methods exist for this conversion, some of which are exact and others iterative. A particularly elegant method attributed to Bowring (1976) is essentially an iterative approach, although accuracy to within about 1 μm can be obtained for terrestrial applications without iteration. For applications far beyond the earth's surface (e.g., satellite applications), a single iteration will yield submillimeter accuracy for points anywhere within the solar system. The applicable equations follow.

(F-6)

(F-7)

To iterate, using the value of ϕ computed by Eq. (F-7), compute an updated value for β by

(F-8)

and use the updated β value in Eq. (F-7) to obtain an updated value for ϕ.

Once the value for latitude has been computed, the ellipsoid height can be computed by

(F-9)

In Eq. (F-9), N must be computed from Eq. (F-4) while using the final value of ϕ. As the latitude approaches the pole, the cosine term in Eq. (F-9) approaches zero, causing the solution for h to be unstable. In that case, Eq. (F-3) can be rearranged to solve for h.

Example F-1

Given the following values of latitude, longitude, and ellipsoid height for a point, and the parameters for the GRS80 ellipsoid, compute the XYZ geocentric coordinates of the point. Check by performing the conversion in reverse.

image

Solution Compute value of normal N by Eq. (F-4).

Compute XYZ coordinates by Eqs. (F-1) through (F-3).

image

Check Compute the value of longitude by Eq. (F-5).

Compute the initial value for ß, using Eq. (F-6).

image

where

Compute the initial value for ϕ, using Eq. (F-7).

image

Since this example was based on a point at the earth's surface, the resulting latitude value should be sufficiently accurate; however, for demonstration purposes an additional iteration will be performed.

Compute an updated value for β, using Eq. (F-8).

Compute an updated value for ϕ, using Eq. (F-7).

image

The final step is to compute the ellipsoid height. Using the value for ϕ computed above, compute the length of the normal N.

Compute the value for h by Eq. (F-9).

F-4 Conversion Between Geocentric and Local Vertical Coordinates

Geocentric and local vertical coordinates are both three-dimensional cartesian coordinate systems. The steps required to convert between these systems involve only rotations and translations. Figure 5-6 shows the relationship between the geocentric XYZ and local vertical $X_l\ Y_l\ Z_l$ systems. Initially, two parameters ϕ_o and X_o are selected which specify the geodetic coordinates of the origin for the local vertical system.

Step 1: Translation to Local Vertical Origin

Beginning with the geocentric XYZ axes with the origin at the center of the earth, a new system $X'Y'Z'$ with its axes parallel to the geocentric axes has been translated by X_o, Y_o, and Z_o such that its origin is at the local vertical system origin. This new system is illustrated in Fig. F-2. Equations (F-10) express the translated coordinates of a point P in the $X'Y'Z'$ system in terms of its $X_P\ Y_P\ Z_P$ geocentric coordinates after the translation.

FIGURE F-2 The $X'Y'Z'$ coordinate system translated to the local vertical origin.

(F-10)

In Eqs. (F-10), X_o, Y_o, and Z_o are the coordinates of the local vertical origin computed from Eqs. (F-1) through (F-3) using the values of ϕ_o and λ_o, and with $h_o = 0$.

Step 2: Rotation of 90° + λ_o about the Z' Axis

Figure F-3 illustrates the geocentric XYZ axes and the $X'Y'Z'$ axes as viewed from positive Z looking toward the earth's equatorial plane. A rotation of 90° + λ_o about the Z' axis is applied to the $X'Y'Z'$ system to create an $X''Y''Z''$ system with its Y'' axis in the direction of local north (i.e., pointing toward the pole in the meridian through the origin). Equations (F-11) express the coordinates of point P in the $X''Y''Z''$ system in terms of X'_P, Y'_P, Z'_P, and the longitude of the local origin λ_o.

(F-11)

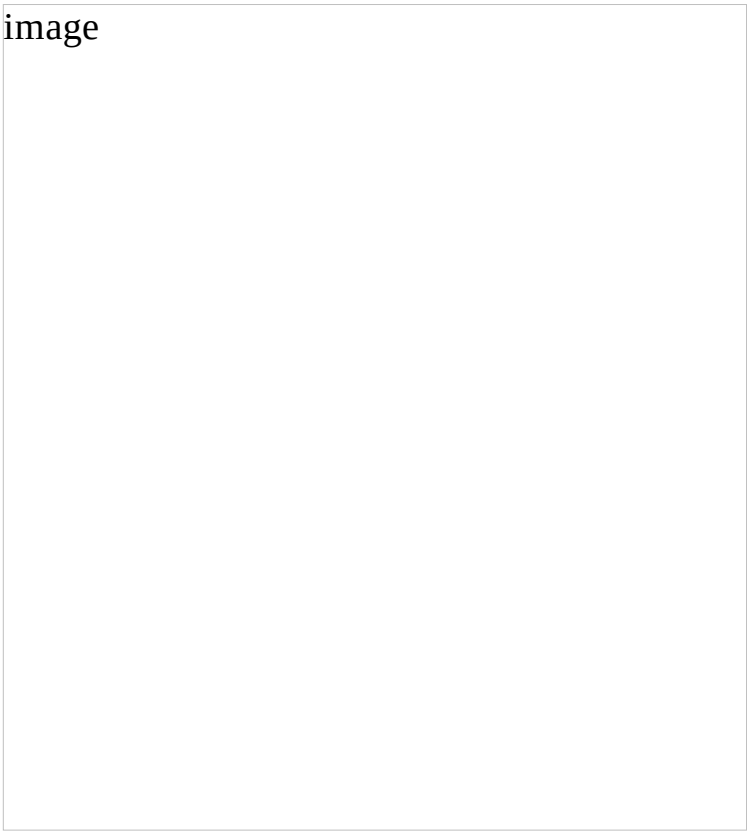


FIGURE F-3 Rotation of 90° + λ_o about the Z' axis.

Step 3: Rotation of 90° – ϕ_o about the X'' Axis

Figure F-4 illustrates the $X''Y''Z''$ system as viewed from the positive X'' axis toward the local vertical origin. A rotation of 90° – ϕ_o will be applied to align the Z'' axis with the normal to the local vertical origin. This gives the final $X_l\ Y_l\ Z_l$ system. Equations (F-12) express the coordinates of point P in the $X_l\ Y_l\ Z_l$ system in terms of X''_P, Y''_P, Z''_P and the latitude of the local vertical origin ϕ_o.

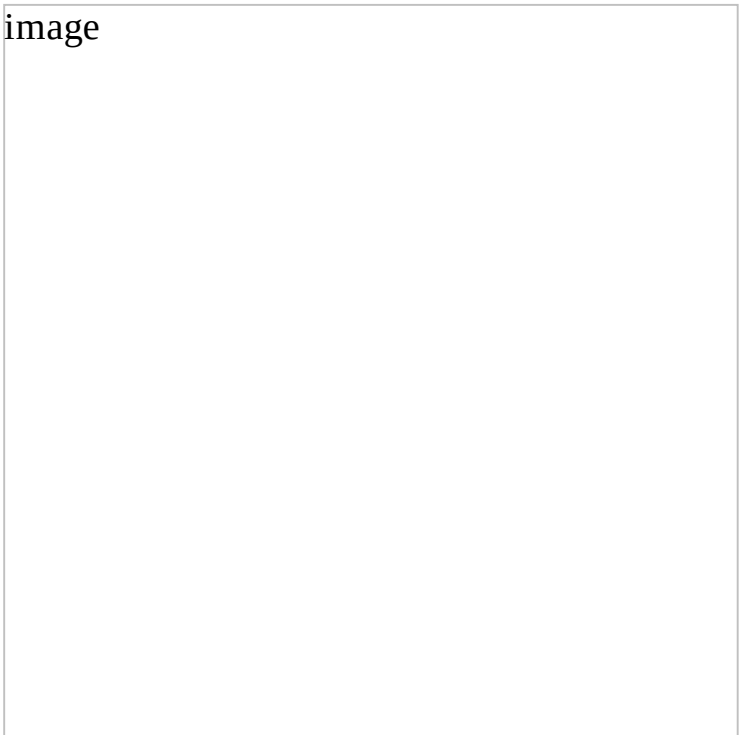


FIGURE F-4 Rotation of $90° - \phi_o$ about the X'' axis.

image

(F-12)

Substituting Eqs. (F-10) into Eqs. (F-11) and in turn substituting the result into Eqs. (F-12), and expressing the result in matrix form, gives

image

(F-13)

Combining the two rotation matrices from Eq. (F-13) into a single rotation matrix M and dropping the P subscripts gives

image

(F-14)

where

image

Equation (F-14) gives the final form of the equation for converting coordinates from geocentric XYZ to local vertical $X_l\ Y_l\ Z_l$. In an optional step, false XYZ offsets can be added as a convenience to cause all coordinates within the project area to be positive.

To convert coordinates of points from local vertical $X_l\ Y_l\ Z_l$ to geocentric XYZ, Eq. (F-14) is rearranged as follows:

(F-15)

Matrix *M* is an orthogonal matrix which has the property that its inverse is equal to its transpose, and therefore Eq. (F-15) can be rewritten as

(F-16)

Equation (F-16) gives the final form of the equation for converting coordinates from local vertical X_l Y_l Z_l to geocentric *XYZ*. If the optional false *XYZ* offsets were previously added, they must be subtracted before Eq. (F-16) is used.

Example F-2

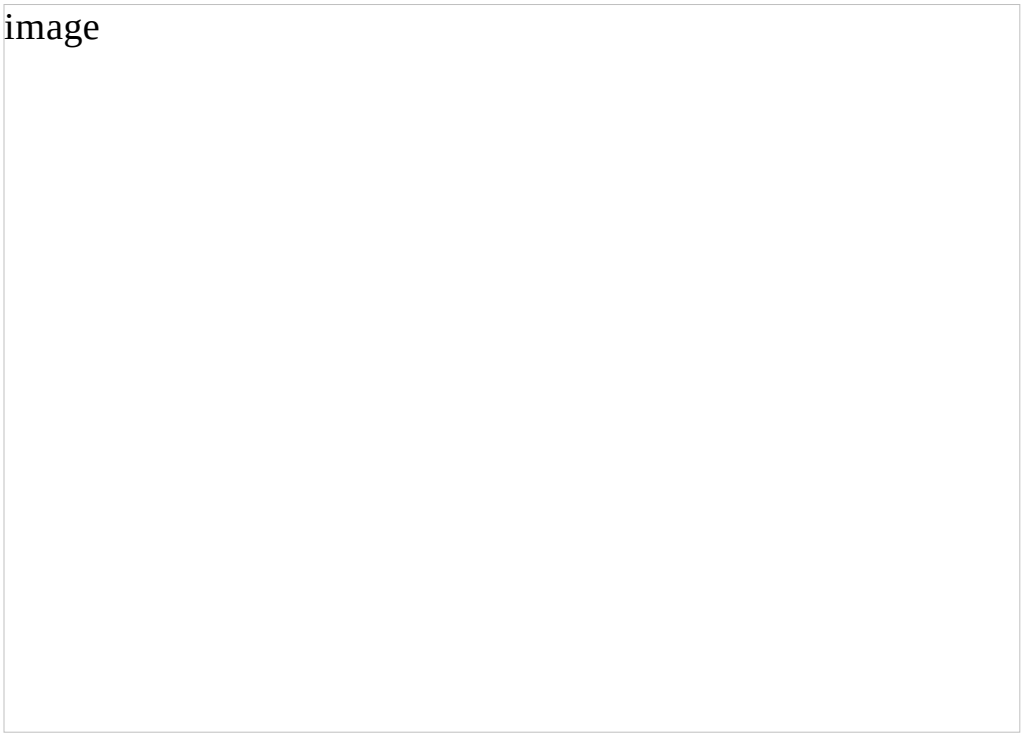
Convert the following *XYZ* geocentric coordinates to local vertical coordinates X_l Y_l Z_l having the parameters listed below. Use ellipsoid parameters for GRS80. Check by performing the conversion in reverse.

Solution

Compute the length of the normal to the local vertical origin by Eq. (F-4).

Compute geocentric coordinates of the local vertical origin by Eqs. (F-1) through (F-3).

Compute rotation matrix terms from Eq. (F-14).



Now solve Eq. (F-14), using the above rotation matrix M, for the local vertical coordinates.

Check Convert back to the original geocentric coordinates, using Eq. (F-16).

F-5 Conversion Between Geodetic and Lambert Conformal Conic Coordinates

Conversion of coordinates between the geodetic system of latitude and longitude, and the Lambert conformal conic map projection, involves complex mathematics which are beyond the scope of this text. The reader may consult references at the end of this appendix for details. The equations presented in this section are from Snyder (1987).

As mentioned in Sec. 5-6, a Lambert conformal conic coordinate zone has a number of defining constants in addition to the required standard ellipsoid parameters. These defining constants are, respectively, ϕ_1 and ϕ_2, the latitudes of the standard parallels; ϕ_o and λ_o, the latitude and longitude of the grid origin; and E_o and N_o, the false easting and false northing.

A forward procedure that can convert latitude and longitude ϕ and λ of a point to X and Y Lambert conformal conic coordinates starts by computing a number of additional parameters which will be constant for a given set of zone constants and ellipsoid parameters. They are

(F-17)

(F-18)

(F-19)

(F-20)

(F-21)

image

After these initial parameters have been computed, the remainder of the forward procedure is as follows. Beginning with the latitude ϕ and longitude λ of a point, compute the following:

(F-22)

(F-23)

(F-24)

(F-25)

An inverse procedure that can convert *XY* Lambert conformal conic coordinates of a point to latitude ϕ and longitude λ requires the same additional parameters that were computed in Eqs. (F-17) through (F-19). Given these parameters, and the *XY* coordinates, the latitude and longitude ϕ and λ can be computed as follows:

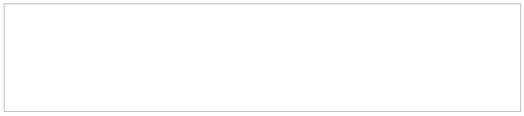
(F-26)

(F-27)

(F-28)

(F-29)

(F-30)

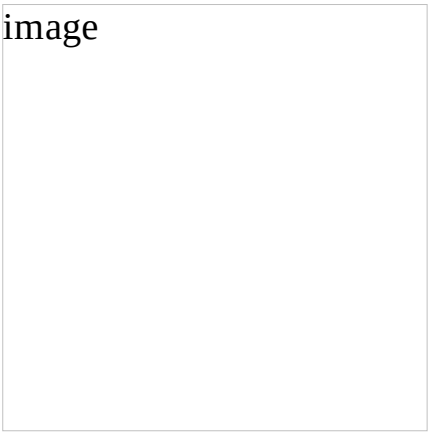


(F-31)

where a = length of semimajor axis
e = first eccentricity

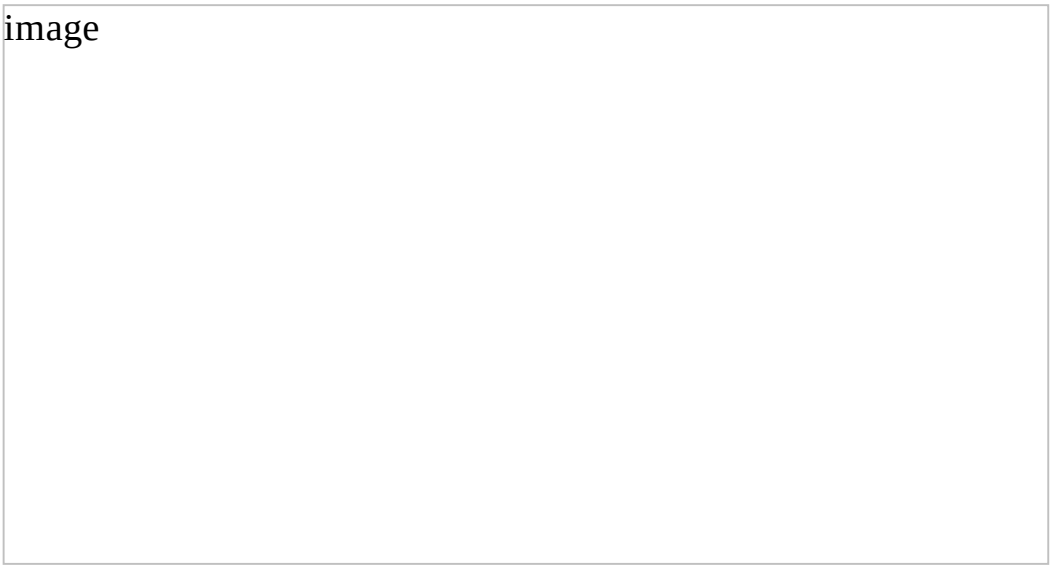
In Eqs. (F-26) and (F-28) the function sign(n) is +1 if n [obtained from Eq. (F-17)] is positive, and it is -1 if n is negative. The inverse tangent function in Eq. (F-28) should be of full-circle range since θ can theoretically range from –180° to +180°, although in practical cases the range will be far less. Equation (F-30) provides an initial approximation for ϕ. This value is then used on the right-hand side of Eq. (F-31) to compute a better approximation for ϕ. This and other successively better approximations can be substituted on the right side of Eq. (F-31) until the value for ϕ stabilizes.

Example F-3
Given the following values ϕ and λ for latitude and longitude of a point, and the specified zone constants for a Lambert conformal conic projection, compute the X and Y coordinates for the point. Use the parameters for the GRS80 ellipsoid. Check by performing the inverse conversion.



Solution Compute required m_ϕ values by Eq. (F-20).

Compute required t_ϕ values by Eq. (F-21).



Compute the additional parameters by Eqs. (F-17) through (F-19).

Compute the t_ϕ value required in Eq. (F-22).

image

Compute values for Eqs. (F-22) through (F-25).

image

Check The check will be performed by inverse conversion, back to ϕ and λ. Compute values for ρ, t, and θ by Eqs. (F-26) through (F-28).

image

image

Compute longitude by Eq. (F-29).

Now for the iterative part. First compute an initial approximation using Eq. (F-30).

Apply Eq. (F-31) to compute updated values for ϕ until it converges.

image

The next iteration produces a negligible change. Therefore,

F-6 Conversion Between Geodetic and Transverse Mercator Coordinates

Conversion of coordinates between the geodetic system of latitude and longitude and the transverse Mercator map projection also involves complex mathematics which are beyond the scope of this text. Again, the reader may consult references at the end of this appendix for details. The equations presented in this section are also from Snyder (1987).

As mentioned in Sec. 5-6, a transverse Mercator coordinate zone has a number of defining constants in addition to the required standard ellipsoid parameters. These defining constants are k_o, the scale factor along the central meridian; ϕ_o and λ_o, the latitude and longitude, respectively, of the grid origin; and E_o and N_o, the false easting and false northing. The longitude of the grid origin λ_o is conventionally referred to as the longitude of the central meridian.

Unlike the Lambert conformal conic conversion equations, those of the transverse Mercator involve series expansions, truncated to a limited number of terms. As such, the accuracy is not limited strictly to the number of significant figures in the computations; instead, it is limited by the truncation of the series. When the following procedures are used, the accuracy should be satisfactory (i.e., subcentimeter) as long as the points are limited to the defined region of the particular transverse Mercator zone. Beyond that, the accuracy deteriorates, particularly in the direction of longitude. The reader is therefore cautioned to use these conversion procedures only within the defined limits of the zone.

A key parameter involved in transverse Mercator conversions is the meridional distance M from

the equator to a specific latitude ϕ. Calculation of M can be performed by a truncated series expansion which is given in the following equation:

image

(F-32)

In Eq. (F-32), a is the semimajor axis, and e the eccentricity. The value of ϕ (latitude) in the first term must be in radians. This equation is accurate to within 1 mm for any latitude.

A forward procedure that can convert latitude ϕ and longitude λ of a point to X and Y transverse Mercator coordinates begins by computing the following preliminary quantities T, C, and A.

(F-33)

(F-34)

(F-35)

where e' = second eccentricity and λ_o = longitude of the grid origin (central meridian). Next, compute the values

image

The following equations complete the forward conversion to X and Y.

(F-36)

image

(F-37)

where k_o = scale factor at central meridian

e' = second eccentricity

E_o, N_o = false easting and false northing, respectively

An inverse procedure that can convert transverse Mercator coordinates X and Y of a point to latitude ϕ and longitude λ begins by computing the *footprint latitude* ϕ_1. The footprint latitude is the latitude at the central meridian that has the same Y coordinate as the point being converted. The footprint latitude is computed by the following equations:

(F-38)

(F-39)

(F-40)

image

(F-41)

In Eq. (F-38), M_o is the meridional distance from the equator to the latitude of the grid origin ϕ_o, as computed by Eq. (F-32); N_o is the false northing; and k_o is the scale factor at the central meridian. The variable μ, in Eqs. (F-39) and (F-41), is the *rectifying latitude* and has units of radians. The units of ϕ_1 in Eq. (F-41) are radians.

Next, compute R_1, the radius of the meridional arc at the footprint latitude.

(F-42)

Then compute the parameters C_1, T_1, and N_1 corresponding to the footprint latitude ϕ_1 by Eqs. (F-34), (F-33), and (F-4), respectively.

image

Now compute the parameter D.

(F-43)

The inverse conversion is then completed by solving the following equations:

image

(F-44)

image

(F-45)

Example F-4
Given the following values for latitude ϕ and longitude λ of a point, and the specified zone constants for the universal transverse Mercator zone 17 projection, compute the X and Y coordinates for the point. Use the parameters for the GRS80 ellipsoid. Check by performing the inverse conversion.

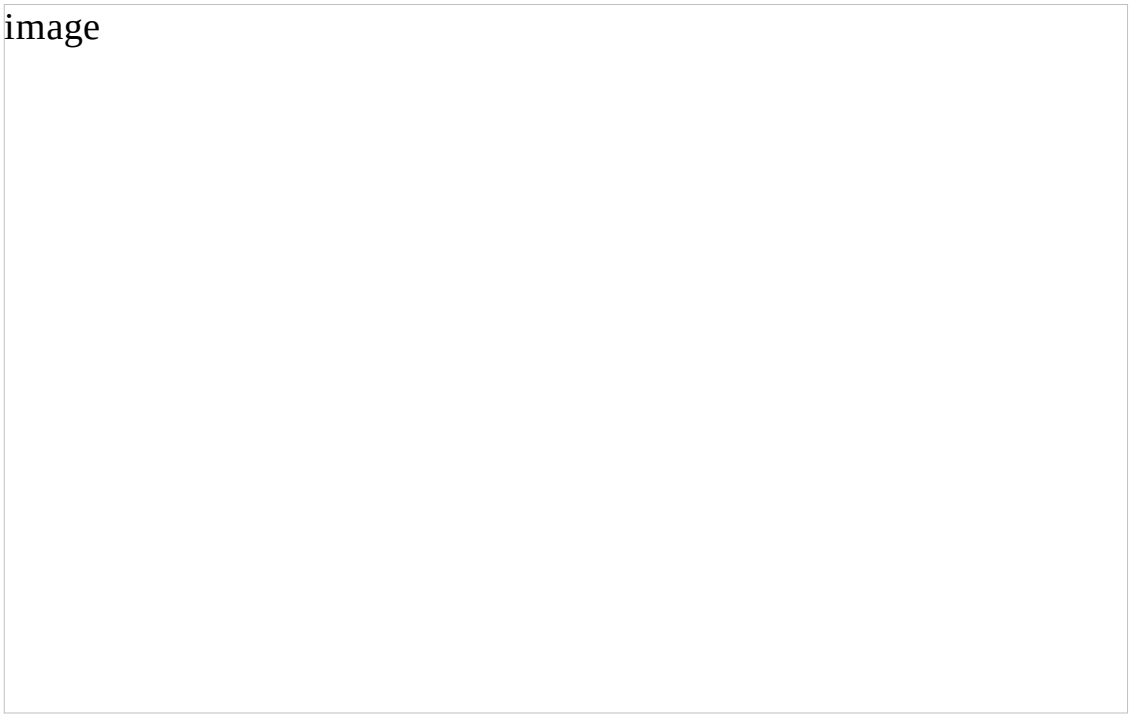
image

Solution Convert ϕ and λ to radians.

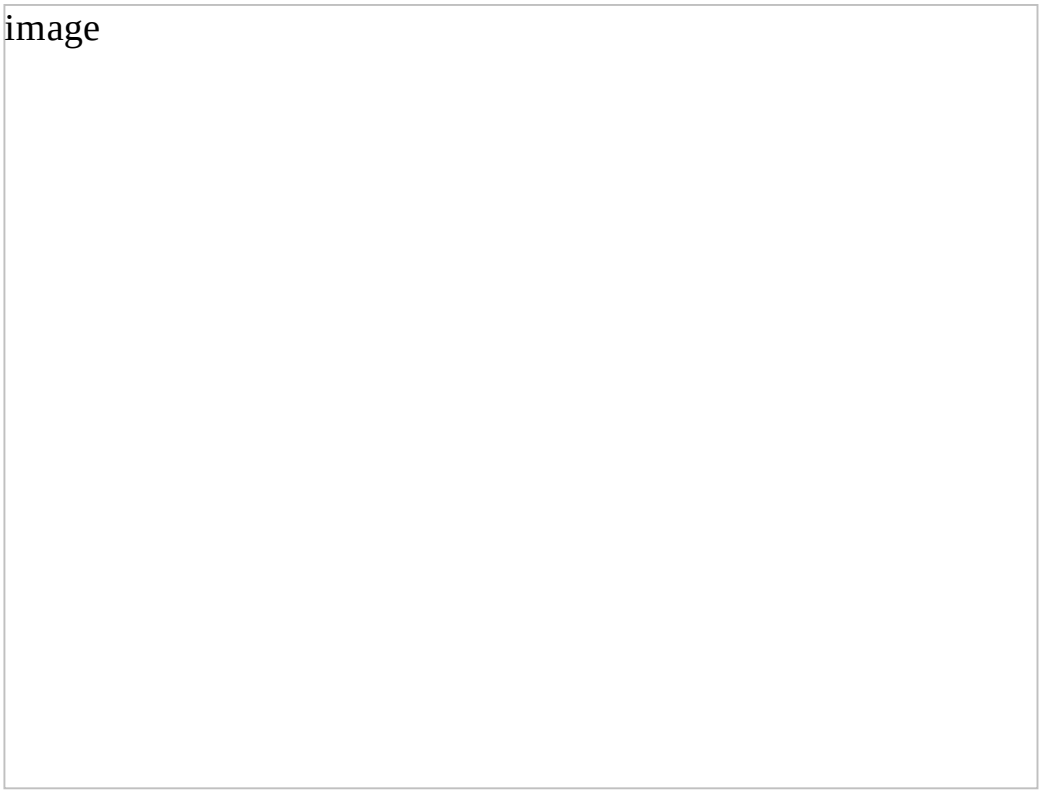
Compute values T, C, and A by Eqs. (F-33) through (F-35).

image

Compute the value of N by Eq. (F-4) and M and M_o by Eq. (F-32).



Now solve Eqs. (F-36) and (F-37) for X and Y.



Check First, compute the footprint latitude by solving Eqs. (F-38) through (F-41).

image

Compute R_1 by Eq. (F-42).

Compute C_1, T_1, and N_1 by Eqs. (F-34), (F-33), and (F-4), respectively.

image

Compute parameter D by Eq. (F-43).

Compute ϕ and λ by Eqs. (F-44) and (F-45).

References

American Society for Photogrammetry and Remote Sensing: *Manual of Photogrammetry*, 5th ed., Bethesda, MD, 2004, chap. 3.

Anderson, J. M., and E. M. Mikhail: *Surveying: Theory and Practice*, WCB/McGraw-Hill, New York, 1998.

Bomford, G.: *Geodesy*, 4th ed., Clarendon Press, Oxford, 1980.

Ewing, C. E., and M. M. Mitchell: *Introduction to Geodesy*, Elsevier, New York, 1970.

Ghilani, C. D., and P. R. Wolf: *Elementary Surveying: An Introduction to Geomatics*, 13th ed., Prentice Hall, NJ, 2012.

Snyder, J. P.: *Map Projections—A Working Manual*, U.S. Geological Survey Professional Paper 1395, U.S. Geological Survey, Washington, 1987.

Stern, J. E.: *State Plane Coordinate System of 1983*, NOAA Manual NOS NGS 5, National Oceanic and Atmospheric Administration, Rockville, MD, 1989.

Problems

F-1. Prepare a computer spreadsheet that will convert geodetic latitude, longitude, and ellipsoid height to geocentric *XYZ* coordinates. Using the spreadsheet, convert the following geodetic coordinates to geocentric. Use the GRS80 ellipsoid.

(*a*) ϕ = 42°06′42.3245″ north, λ = 93°33′45.2165″ west, h = 276.409 m
(*b*) ϕ = 37°03′41.2227″ north, λ = 122°32′56.1093″ west, h = 153.131 m
(*c*) ϕ = 27°59′26.9389″ north, λ = 86°56′31.0057″ east, h = 8847.734 m
(*d*) ϕ = 64°24′30.1990″ north, λ = 152°02′43.6737″ west, h = 6193.536 m

F-2. Prepare a computer spreadsheet that will convert geocentric *XYZ* coordinates to geodetic latitude, longitude, and ellipsoid height. Using the spreadsheet, convert the following geocentric coordinates to geodetic. Use the GRS80 ellipsoid.

(*a*) X = –1,777,460.052 m, Y = 2,921,841.730 m, Z = 5,370,754.340 m
(*b*) X = 855,525.030 m, Y = –5,487,931.094 m, Z = 3,125,080.292 m
(*c*) X = –4,592,671.234 m, Y = 2,537,259.009 m, Z = –3,617,548.719 m
(*d*) X = 4,054,206.962 m, Y = 2996.527 m, Z = 4,907,448.862 m

F-3. Prepare a computer spreadsheet that will convert geocentric *XYZ* coordinates to local vertical coordinates X_l Y_l Z_l. Using the spreadsheet, convert the following geocentric coordinates to local vertical. Use the GRS80 ellipsoid and a local vertical origin latitude ϕ_o = 29°30′00″ north and longitude λ_o = 82°12′00″ west.

(*a*) X = 780,656.787 m, Y = -5,501,656.889 m, Z = 3,120,519.196 m
(*b*) X = 778,970.096 m, Y = -5,453,583.214 m, Z = 3,203,737.478 m
(*c*) X = 720,420.933 m, Y = -5,481,878.690 m, Z = 3,169,160.707 m
(*d*) X = 734,256.533 m, Y = -5,527,174.279 m, Z = 3,086,742.792 m

F-4. Prepare a computer spreadsheet that will convert local vertical coordinates X_l Y_l Z_l to geocentric *XYZ* coordinates. Using the spreadsheet, convert the following local vertical coordinates to geocentric. Use the GRS80 ellipsoid and a local vertical origin latitude ϕ_o = 29°24′00″ north and longitude λ_o = 81°50′00″ west.

(*a*) X_l = –18,738.640 m, Y_l = 14,769.927 m, Z_l = 1.981 m
(*b*) X_l = 7760.644 m, Y_l = 2268.638 m, Z_l = 45.884 m
(*c*) X_l = 5171.700 m, Y_l = –6029.912 m, Z_l = 43.388 m
(*d*) X_l = –25,068.884 m, Y_l = –20,825.035 m, Z_l = –30.919 m

F-5. Prepare a computer spreadsheet that will convert geodetic latitude and longitude to Lambert conformal conic *XY* coordinates. Using the spreadsheet, convert the following geodetic coordinates to Lambert conformal conic. Use the GRS80 ellipsoid and the following Lambert zone constants: ϕ_1 = 42°44′ north, ϕ_2 = 44°04′ north, ϕ_o = 42°00′ north, λ_o = 90°00′ west, E_o = 600,000 m, and N_o = 0.

(*a*) ϕ = 42°45′23.5233″ north, λ = 91°14′59.0517″ west
(*b*) ϕ = 44°23′51.3480″ north, λ = 90°09′20.2099″ west
(*c*) ϕ = 43°43′42.6206″ north, λ = 89°24′38.0503″ west
(*d*) ϕ = 44°27′15.0788″ north, λ = 91°21′26.7792″ west

F-6. Prepare a computer spreadsheet that will convert Lambert conformal conic *XY* coordinates to

geodetic latitude and longitude. Using the spreadsheet, convert the following Lambert conformal conic coordinates to geodetic. Use the GRS80 ellipsoid and the following Lambert zone constants: ϕ_1 = 45°34′ north, ϕ_2 = 46°46′ north, ϕ_o = 45°10′ north, λ_o = 90°00′ west, E_o = 600,000 m, and N_o = 0.

(*a*) X = 498,129.633 m, Y = 142,242.591 m
(*b*) X = 565,869.445 m, Y = 85,860.635 m
(*c*) X = 616,953.581 m, Y = 40,639.790 m
(*d*) X = 534,976.172 m, Y = 184,256.842 m

F-7. Prepare a computer spreadsheet that will convert geodetic latitude and longitude to transverse Mercator *XY* coordinates. Using the spreadsheet, convert the following geodetic coordinates to transverse Mercator. Use the GRS80 ellipsoid and the following transverse Mercator zone constants: k_o = 0.9999411765, ϕ_o = 24°20′ north, λ_o = 81°00′ west, E_o = 200,000 m, and N_o = 0.

(*a*) ϕ = 28°37′39.5323″ north, λ = 82°02′35.0368″ west
(*b*) ϕ = 26°38′29.5273″ north, λ = 81°49′37.0588″ west
(*c*) ϕ = 27°25′37.3755″ north, λ = 81°46′40.8965″ west
(*d*) ϕ = 25°18′20.9991″ north, λ = 80°44′29.2206″ west

F-8. Prepare a computer spreadsheet that will convert universal transverse Mercator (UTM) *XY* coordinates to geodetic latitude and longitude. Using the spreadsheet, convert the following UTM coordinates to geodetic. Use the GRS80 ellipsoid and the following UTM zone constants: zone 17 k_o = 0.9996, ϕ_o = 0°00′, λ_o = 81°00′ west, E_o = 500,000 m, and N_o = 0.

(*a*) X = 502,841.680 m, Y = 3,588,973.376 m
(*b*) X = 394,120.551 m, Y = 3,257,138.082 m
(*c*) X = 465,262.738 m, Y = 2,801,506.948 m
(*d*) X = 367,112.167 m, Y = 3,016,226.144 m

A

A posteriori standard deviation
A priori standard deviation
Aberrations
Abscissa scale
Absolute orientation
    analytical plotters
    direct optical projection stereoplotters
Accelerometers, usage
Accidental errors
Accommodation
Accuracy, concepts
Acidic solution
ACSM. *See* American Congress on Surveying and Mapping
Adjacent flight strips, photography
Adjusted quantities
    precisions, matrix equations
    standard deviation
ADS80 airborne digital sensor (Leica)
Aerial film speed (AFS)
Aerial image, wavelet decomposition level
Aerial mapping:
    cameras, filters (usage)
    metric cameras, usage
Aerial mosaics:
    classes
    photographs, usage
Aerial photogrammetry, control
Aerial photographic mission
Aerial photographs:
    camera orientation
    examples
    overlapping pair
    stereopair, space intersection
    tilted aerial photographs, atmospheric refraction
    usage
Aerial photography
    acquisition
    atmospheric refraction
    availability
    classification
    crab, presence
    depth of field, impact
    exposure

laser scan data, combination
overlapping aerial photography, geometry
photogrammetrist/photo interpreter collection
planning
neat model, usage (transparent template)
stereoscopic plotting instrument design
sun angle
vertical photos, calculations
Aerotriangulation
airborne linear array sensors
analytical plotters, application
computational strategies
image points, photos (example)
overlapped position, photos (block)
pass points
locations
performing, methods (classification)
planning
reduced equation (structure)
cross-strip ordering, usage
down-strip ordering, usage
usage
value
Affine coordinate transformation
Affine transformation:
application
computation
linear form, obtaining
solving
usage
AFS. *See* Aerial film speed
AGPS. See Airborne global positioning system
Air base
computation
Air photo interpretation, employment
Air turbulence, impact
Airborne control
usage
Airborne global positioning system (AGPS) systems, deployment
Airborne GPS control:
bundle adjustment
camera/IMU/GPS antenna, configuration
flight strips, configuration
Airborne GPS positioning, considerations
Airborne GPS-INS system
Airborne laser scanning (ALS)
Optech Orion airborne laser scanning system

swath pattern
system, swath pattern
unfiltered/filtered ALS data
Airborne linear array sensors:
distortion, air turbulence (impact)
usage
Aircraft vibrations, detrimental effects
Air-to-glass interface
Alignment, correction
ALS. *See* Airborne laser scanning
Altimeters, usage
Altitude (h)
American Congress on Surveying and Mapping (ACSM)
American Society for Photogrammetry and Remote Sensing (ASPRS)
founding
photogrammetry definition
American Society of Civil Engineers (ASCE)
Geomatics Division
Amplitude
Anaglyphic system
complementary color filters, usage
Analog form (continuoustone image)
Analog plotters:
absolute orientation
analytical plotters, comparison
Analytic relative orientation
Analytical absolute orientation
Analytical aerotriangulation techniques, variations (evolution)
Analytical interior orientation
Analytical photogrammetry
computations
Analytical plotters
absolute orientation
analog plotters, comparison
application
capabilities
comparator, schematic representation
components, schematic diagram
interior orientation
mathematical model, computation
operation
method
schematic diagram
operation, profiling mode
orientation
P-3 analytical plotter (Zeiss)
real-time operation, top-down flow

relative orientation
reset capability
SD3000 analytical plotter (Leica)
system components
three-dimensional operation
usage
modes
vector superimposition
Analytical rectification
Analytical relative orientation
collinearity, usage
stereopair, example
Analytical self-calibration
collinearity equations, usage
example
matrix equations
numerical stability, concern
usage
Analytical stereomodel
formation
Analytical stereoplotters
CAD computer
encoders
grid plate, placement
half mark movement, approaches
operator controls
principal distance, setoff
servomotors
three-dimensional operation, flowchart
Angle equivalents
Angle of incidence
Angle of refraction
Angular distortion
Angular field of view
Angular orientation
chaining
expression
omega-phi-kappa
Antenna-to-camera vector
Aperture:
diameter
decrease
increase
priority mode
shutter speed, relationship
Arago, Francois
Area equivalents

Area-based digital image-matching techniques
Area-weighted average resolution (AWAR)
Artificial photogrammetric targets
Artificial targets
    disadvantages
    example
    usage
ASCE. *See* American Society of Civil Engineers
ASCII editor, usage
ASCII text editor, usage
ASPRS. *See* American Society for Photogrammetry and Remote Sensing
Asymmetric radial lens distortion
ATAN2
Atmospheric haze:
    cause
    scatter
Atmospheric refraction:
    aerial photography
    correction
        computation procedure
    distortions
        modeling, inadequacy
    impact
Atmospheric scatter
AutoBar exterior orientation device
Autocollimation, principal point
Automatic data recording
Auxiliary coordinate system, point requirement
Auxiliary tilted photo coordinate system
Auxiliary x'y' image coordinate system
Average photo scale
Average terrain, map scales/contour intervals
AWAR. *See* Area-weighted average resolution
Azimuth (a)
    angular orientation
    calculation, normal vectors (usage)
    computation
    differences
    rotation
        angles
        example
Azimuth-tilt-swing rotation, usage

B

Backing, photographic film component
Back-surfaced mirror
Backward linear array sensor scans

Bandpass filter
Bandpass image
Base-classes
Base-height (B/H) ratio
    increase
Baseline vectors, computation
Basic control
b_e/h. *See* Stereoviewing
Best symmetry, point
ß angles, algebraic signs
Between-the-lens shutters
BGR. *See* Blue Green Red
B/H. *See* Base-height
Bicubic interpolation
    discussion
    shape
Bilinear interpolation
    advantage
    shape
Bilinear/bicubic interpolation, usage
Binary digits (bits)
Binary images
Binocular attachment
Black-and-white emulsions, sensitivities
Black-and-white infrared photographs
Black-and-white photographs, processing/printing
    steps
Blade-type lens shutters
Block-diagonal matrix, nonzero submatrices
Block-diagonal submatrix
Blue Green Red (BGR):
    color cube
    color space
    coordinates, conversion
        example
Blunders
    elimination
Blurring
Boresight angles
Boresight calibration, facilitation
Boresight rotation matrix
Bridging
    usage
Brightness factor
Bundle adjustment. *See* also Simultaneous bundle adjustment
    cause
    convergence

failure
initial approximations
measurements
post-bundle adjustment statistics
quantities, obtaining
results, interpretation
self-calibrating bundle adjustment
Bundle of rays
Bureau of Reclamation
Bytes

C

C factor
C/A. *See* Coarse acquisition
Calibrated focal length (CFL)
basis
Calibrated principal point
Calibration parameters:
loosening, adjustment divergence
weighted observations
Calibration target
template
Camera calibration
collimators, usage
field methods
goniometer laboratory procedure
laboratory methods
parameters
loosening (examples)
resolution test pattern
stellar methods
Cameras
angular field of view
axis, inclination angle
determination
variable, importance
body
configuration
controls
exposure stations, angular orientation (impact)
fiducial marks
focal length, knowledge
interior orientation
lens
distortion
incident nodal point
magazine

Manual of Photogrammetry definition
mounts
orientation, aerial photographs
resolution, calibration
single-lens camera
Camera-to-antenna vector
Camouflage detection
Canadian Institute of Geomatics (CIG)
Carrier-phase processing, usage
CCD. *See* Charge-coupled device
Central meridian
CFL. *See* Calibrated focal length
Chain method, usage
Chaining
Characteristic curve. *See* also Photographic emulsions
usefulness
Charge-coupled device (CCD)
array
distortions
array/film
elements
frame, width
solid-state CCD imaging array
Checkpoints, usage
Chip rate
Chi-squared test
CIG. *See* Canadian Institute of Geomatics
Clarke 1866 ellipsoid, usage
Close-range photogrammetry
accuracy
applications
examples, parabolic antenna shape
object depth, importance
planning
ship propeller size/shape, determination
Close-range photography:
camera exposure stations, angular orientation (impact)
geometric consideration
planning, physical constraints
Close-range stereo coverage
CMOS. *See* Complimentary Metal-Oxide Semiconductor
Coarse acquisition (C/A):
code
tracking
Code-phase processing
Collimator targets, images
Collinearity

applications
modified collinearity equations, usage
space intersection
space resection
usage
Collinearity condition
equations, development
examples
expression
Collinearity equations
azimuth-tilt-swing rotation, usage
homogeneous representation
linearization
nonlinear characteristics
quantities, determination
usage
Color additive process
Color cube:
blue-green-red color cube
viewpoints
Color exposure, making
Color film
cross section
layers, color sensitivity
Color image:
example
representation
Color infrared emulsions, usage
Color infrared film (false-color film):
development
sensitivity
Color infrared image
Color infrared photographs
Color negative (color reversal film)
Color reversal film (color negative)
Color subtractive process
Color values, basis
Column/row image coordinates
Combination lens
Common line, swing (calculation)
Community rating system (CRS)
Comparator:
monocomparators/stereocomparators, usage
schematic representation
two-dimensional affine relationship
Compensating errors
Complimentary Metal-Oxide Semiconductor (CMOS) devices

Composite orthophoto
Computer-aided drafting (CAD)
  CAD-unit measurements
  computer, usage
  drawings
  environment
  program
  systems
    analytical plotter interface
Computerized mapping systems, usage
Conformal applications
Conformal coordinate transformation
Conformal transformation
Conjugate principal points
  distances
Contact printer
Contact printing
Contact prints
Contours:
  direct compilation, stereoplotter (usage)
  direct tracing
  generation
  intervals
    factors
  locations, inconsistencies
  maps
    terrain, representation
  renditions
Contrast enhancement
Control configuration
Control points
  absence
  ground/image coordinates,
  horizontal control points
  images, blurred edges
  location
  number, minimum (usage)
  stereomodel coordinates
  vertical control points
Controlled aerial mosaics
Controlled digital mosaics
Converging lens
  focal length
Convex glass lens, light entry
Convolution
  cubic convolution
  kernel

mathematical operation
performing
Coordinate transformations
application
approach
computation
method, alternative
working sketches, usage
effects
matrix methods
redundancy, impact
Coordinates:
computation
correction, computation
determination
reduction
systems
transformation, preparation
Coplanarity
equation
linearization
writing
usage
Coplanarity condition
equation
development
usage
vectors, usage
example
Cornea
Correction plate, usage
Correlated, term (usage)
Correlation
Correlation coefficients
computation
moving windows, usage
selection
summation terms, computation
Crab
correction
Crabbed overlapping aerial photographs
Cross-strip ordering
strategy
usage
CRS. *See* Community rating system
Cubic convolution
Cubic spline approximation

Cylindrical intensity-huesaturation system, gray line position
Cylindroid

D

Daguerre, Louis
Darkness, degree
Data block
Data merging
Data reduction techniques
Datum nadir point
Datum plane
Datum principal line
Datum principal point
Dead reckoning
Decentering (symmetrical radial distortion)
    computation
    pattern
Decentering lens distortion
Deconvolution
Deeply coupled integration level
Degrees of freedom, definition
Delaunay triangulation
DEMs. *See* Digital elevation models
Densitometer, usage
Density
Depression angle
Depressions, appearance
Depth of field
    direct optical projection stereoplotters
Depth perception
    relationships
Derivative terms, interpretation
Developable surface
Developer
Developing, black-and-white photographic processing/printing step
Deville, Eduard
DFT. *See* Discrete Fourier transform
Diaphragm
    area
        calculation
    diameters
    light regulation
Diapositives (transparencies)
    centering
    direct optical projection stereoplotters
    preparation
Differential calculus, usage

Differential leveling
Differential positioning
Differential rectification (orthorectification)
Differential rotation, measurement
Differential tilt rotations
Digital CAD file, saving
Digital color image, display
Digital elevation models (DEMs)
    accuracy, effect
    automatic production
    automatically matched points, positions
    collinearity relationship
    compilation
    points
        matching
        positions
    rectangular DEM, three-dimensional representation
    usage
Digital form (pixel array)
Digital image (digital images)
    display
    example
    linear stretch contrast enhancement
    matching
    model
    processing, types
    production
    quantization levels, variation
    resampling, example
    spatial frequency
    spatial resolutions, variation (example)
Digital image-matching techniques, categories
Digital imagery, georeferencing
Digital imaging process, principles
Digital line graphs (DLGs), usage
Digital mapping:
    cameras
        formats
    systems, existence
    topographic features, representation
Digital mosaics:
    controlled digital mosaics
    semicontrolled digital mosaics
    uncontrolled digital mosaics
Digital number (DN)
    brightness, correspondence
    computation

histogram, examples
matrix
representation
Digital orthophoto production process, schematic representation
Digital photogrammetry
Digital photographs:
coordinates
joining, tie points (usage)
rectification
Digital printing, color-subtractive process
Digital rectification
equipment, requirement
usage, example
Digital resampling
Digital satellite image, Fourier transform (example)
Digital single-lens reflex camera, example
Digital surface model (DSM)
Digital terrain model (DTM)
generation
Digital-frame cameras:
classification
image subsets, usage
Dimensionless representative fractions/ratios
Direct compilation:
planimetric details
stereoplotter, usage
Direct contact printing
Direct conversion
Direct optical projection instruments
Direct optical projection stereoplotters:
absolute orientation
components
depth of field
diapositives
guide rods
illumination lamps
interior orientation
kappa
leveling screws
main frame
model
scale/base, change/adjustment
tilt, correction
omega
phi
platen
principal components

principal distance
projection distance
projection systems
projector bar
projectors
reference table
relative angular relationship, recreation
relative orientation
tracing pencil
tracing table
viewing/tracing systems
Direction cosines
Discrete Fourier transform (DFT)
Discrete waveform laser scanning system
Displaced point, coordinates (determination)
Distribution curve
DLGs. *See* Digital line graphs
D-log E curve
DN. *See* Digital number
Down-strip ordering
strategy
usage
Drainage, features
Drift
Drying, black-and-white photographic processing/printing step
DSM. *See* Digital surface model
DTM. *See* Digital terrain model
Dual-frequency receivers, usage
Dynamic terrestrial photography
Dynamo close-range camera
strobe attachment

E

Earth curvature:
correction
distortion
Easel plane
Edges:
detection
operator
effects, indication
information
measurement
multiscale representations
EFL. *See* Equivalent focal length
Electromagnetic energy, velocity
Electromagnetic spectrum, wavelength classification

Elements of exterior orientation (EOPs)
Elevation angle
Elevation layers, hydrography layers (combination)
Elevation model (preparation), DEM (usage)
Ellipsoid height:
    computation
    values
Emergent nodal point
Emulsions. *See* also Photographic emulsions
    photographic film component
E'N coordinates
EN ground coordinate system, scaled X'Y' coordinate system (superimposition)
EN system
Encoders, usage
End lap (flight strip photograph). *See* also Percentage of end lap; Photographic end lap
    definition
    overlap area
    photo overlap
Engineering projects, photomaps (advantages)
English engineer's scales
Enlargement:
    geometry
    ratios
Enlargement ratio
Environmental effects, groundwater resource monitoring
Environmental monitoring, photogrammetric applications
Environmental Protection Agency (EPA)
EOPs. *See* Elements of exterior orientation
EPA. *See* Environmental Protection Agency
Epipolar geometry:
    principle, usage
    softcopy plotters
    stereopair
Epipolar plane
Equally weighted observations
Equations, systematic formulation
Equivalent focal length (EFL)
Equivalent length
EROS Data Center, photo archive
Errors. *See* also Systematic errors
    definition
    evaluation
    propagation
Estimated digit
Exposure:
    control
    instant

stations
photographs
variation
Exterior orientation elements
determination
Exterior orientation parameters
approximations
measurements
systemic variations
variation
weight matrix
Eye base
Eye base to perceived model height ratio

F

Falloff
False northing
False-color film (color infrared film):
development
sensitivity
Fast Fourier transform (FFT)
Feature-based digital
image-matching
techniques
Federal Emergency Management Agency (FEMA), floodplain management
Federal Geographic Data Committee
Federal Highway Administration
FFT. *See* Fast Fourier transform
Fiducial axis, principal point (coordinates)
Fiducial coordinates, computation
Fiducial marks
absence
calibrated coordinates
comparator-measured coordinates
coordinates
Field surveying:
economic advantage
usage
Field surveys:
methods, usage
usage
Field-surveyed control, accuracy
Field-surveyed ground control
Film. *See* Photographic film
Filtered ALS data
Filters
purposes

usage
First-level wavelet decomposition
First-surface mirror
touching, avoidance
500 ELX camera (Hasselblad)
Fixing, black-and-white photographic processing/printing step
Flight and Sensor Control Management System
Flight height, planimetric accuracy
Flight lines:
number
spacing
adjustment
Flight map
Flight path, orientation fixation
Flight plan
example
preparation, computer (usage)
project area
Flight planning:
importance
template method
Flight strips
configuration. *See* Airborne GPS control.
overlap
photographs
end lap
side lap
Flight-line axis:
system
usage
Flight-line photographic coordinates, image expression
Floating line
Floating mark
placement
principle
example
usage
Floodplain rating:
example
photogrammetric applications
Flying height
aerial photography, usage
computation
determination
increase, relief displacement (reduction)
measurement
scale variations

variation
impact
Flying spot scanners
FMC. *See* Forward-motion compensation
Focal length (lens). *See* also Calibrated focal length
Focal plane
flatness
shutters
Focal point
Focusable lenses, stability
Focusing, differences
Fog
Footprint latitude
Forward conversion
Forward linear array sensor scans
Forward procedure, impact
Forward-motion compensation (FMC)
Fourier transform:
discrete Fourier transform (DFT)
examples
fast Fourier transform (FFT)
usage
Fractional location, image pixels (4×4 subarea)
Frame aerial cameras:
generalized cross section
main parts
Frame cameras (frame sensors)
single-lens frame cameras
French Academy of Science
Frequency
domain
From system
Front nodal point
F-stop (f-number)
occurrence
setting
variation
shutter speed, relationship
Full error propagation, assumption
Full strip, rotations
Full-frame sensor
Full-waveform laser scanning system
Fully automatic mode

G

Gauss, Karl
Gaussian distribution

Gaussian kernel, usage
GCPs. *See* Ground control points
Geocentric coordinates
    conversion
    geodetic coordinate conversion
    local vertical coordinate conversion
Geodesy, concepts
Geodetic coordinates
    conversion
    Lambert conformal conic coordinates, conversion
    system
    transverse Mercator coordinates, conversion
Geodetic reference surfaces
Geodetic Reference System of 1980 (GRS80)
    ellipsoid constants
    ellipsoid parameters, usage
Geodetic-grade receivers, usage
Geographic information systems (GIS):
    analytical tool
    data, compilation
    development
    information layers, compilation
    objectives
    photogrammetric applications
    photogrammetry, relationship
    planimetric mapping, elementary methods
    relational database
    site development
    software packages
    usage
Geographic information systems (GIS) database:
    digital maps
    information
        complication
        layers, development
    modification
Geomatica
Geomatics Division (ASCE)
Geometric optics
    concept, point source (ray emanation)
Geometric resolution
Geometric strength
Georeferenced images
    assembly
    geometric accuracy
    schematic representation
Georeferencing

application
image alignment
process
Geospatial Positioning Accuracy Standards
GIS. *See* Geographic information systems
Global navigation satellite system (GNSS) signals
Global Positioning System (GPS)
algorithms, emergence
antenna
coordinates, translation
geometric relationship
antenna, configuration
control, bundle adjustment. *See* Airborne GPS control.
differential positioning
fixes, interpolation
fundamentals
kinematic GPS positioning
methods, usage
onboard GPS-INS system
receiver
survey
usage, increase
GLONASS
Goniometer
GPS–INS data, processing
GPS-INS integration
GPS-INS system. *See* also Airborne GPS-INS system
accuracy/reliability
GPS/INS system
Grains
Graphical analysis
Gray level, quantification
Gray line
Grid DEMS, disadvantage
Grid plate:
calibration
placement
Ground control:
coordinates
equations
observations
errors
establishment
existence
images
planning
Ground control points (GCPs)

relief, correction
Ground coordinates
axis system
calculation
conversion
groundel centers, association
system
strip model, adjustment
Ground coverage
square, dimension
Ground distance, measurement
Ground elements (groundels)
Ground EN two-dimensional system
Ground nadir point
Ground registration
Ground system, coordinates
Ground-based system, three-dimensional coordinates
Groundwater resource area, monitoring
GRS80. See Geodetic Reference System of 1980
Guide rods (direct optical projection stereoplotters)
Gyroscopes (gyros):
drift
usage

H

Habitat layer, usage
Half marks
movement, approaches
spacing
H-and-D curve
Hard-copy photographs, usage
HARNs. See High-accuracy reference networks
Hasselblad 500 ELX camera
Heads-up digitizing
Height, isolation
Heliography
Hertz, wavelength
Hidden objects
High oblique images
High oblique photograph:
components
example
High-accuracy reference networks (HARNs)
High-altitude photographs, usage
High-altitude photography, ground coverage per photo
High-frequency information, loss
High-order rational polynomial transformations, usage

High-pass filter, application
High-pass operation
High-resolution digital sensors
Highway planning/design
Hinge point
Histogram
  equalization
    contrast enhancement, digital image
    example
  examples
  graphic display
  production
  shape, variation
Homogeneous coordinates:
  representation, three-dimensional line (usage)
  usage
Horizon, location (oblique photograph)
Horizontal accuracy
Horizontal angles:
  computation, oblique photographs
  measurement
Horizontal control
  establishment, field survey methods (usage)
  images
  points
    requirement
  surveys
Horizontal datums
  theoretical development
Horizontal plane, construction
Horizontal terrestrial photos, obtaining
Hough transform
Housing, gyroscopic stabilization
Hue:
  polar coordinates
  representation
    color value basis
Human eye
  cross section
  eye base
  focusing, differences
  intensity response
  lens, biconvexity
  object, appearance
Hybrid digital image-matching techniques
Hydrologic area units

I

ICP. *See* Iterative closest point
IHS. *See* Intensity-huesaturation
IKONOS:
    QuickBird, contrast
    satellite imagery
Illuminance
    decrease
    distance, inverse proportion
    image distance, impact
    increase
    proportion
    variation
Image pixels, 4×4 subarea
Image-matching techniques, usage
Images. *See* also Binary images; Digital images
    acquisition
        process
    alignment
    approximations, discovery
    classification
    coarseness
    color images
        representation
    color infrared image
    column/row coordinates
    coordinates
        system, rotation
    data, nonimage data (combination)
    digital image display
    digital rectification, usage (example)
    distance
        impact
    distortions, compensation
    enhancement
    formation (single-lens camera)
    interpretation, branches
    measured image coordinates, refinement
    measurements
        coordinate systems
        softcopy plotters
    one-point perspective image
    orientation
    plane
        distortion
    pyramid, formation
    quality, advantages

resampling
sharpness, degradation
spacing, change
systematic degradations, cumulative effect (representation)
three-point perspective image
two-point perspective image
Imaging device
IMU. *See* Inertial measurement unit
INCA3 digital close-range camera
Incidence, angle
Incident light ray:
pencil
refraction
Incident nodal point
Inclination angle, determination
Independent model
Indicated principal point
Indirect contouring
Inertial measurement device
Inertial measurement unit (IMU)
configuration
geometric relationship
orientation, boresight angles (usage)
output
Inertial navigation system (INS)
capability
mechanization
usage
navigation, dead reckoning
Inertial reference frame
Inertial-based navigation, basis
Infinite focus, plane
Inflatable antenna, size/shape (determination)
In-flight camera calibration
Information, inclusion
Infrared (IR) emulsions
Inner cone, usage
INS. *See* Inertial navigation system
Instant of exposure
Instrumentation:
advancements
developments
Integer ambiguity
Integration levels
Intelligent Digital Camera 3 (INCA3)
Intensity-hue-saturation (IHS) system
conversion

Interference noise, Fourier transform
Interior orientation
    analytical interior orientation
    analytical plotters
    direct optical projection stereoplotters
    elements
    parameters, inclusion
International Earth Rotation Service
International Society for Photogrammetry and Remote Sensing (ISPRS), founding
International standard foot
International Standards Organization (ISO) number, usage
International Terrestrial Reference Framework (ITRF)
Interpolation:
    bicubic interpolation
    bilinear interpolation
    row interpolation weight matrix, computation
Interpretative photogrammetry
Intersect program, running
Intervalometer
Inverse distance weighting
Inward radial distortion
ISPRS. *See* International Society for Photogrammetry and Remote Sensing
Iterative closest point (ICP) algorithm
ITRF. *See* International Terrestrial Reference Framework

J

Journal of Spatial Science

K

K matrix, WC contributions
K value, computation
Kalman filter
Kappa
    angle, z_2 axis (rotation)
    direct optical projection stereoplotters
    rotation
Kernels
    application
    Gaussian kernel, usage
    matrix
    values
Kinematic GPS positioning
Kinematic positioning technique
Kriging

L

Lambert, J. H.

Lambert conformal conic conversion equations
Lambert conformal conic coordinates, conversion
Lambert conformal conic projection
Land management, photogrammetric applications
Land use
  layer, land cover layer (combination)
  planning, photomaps (usage)
Laplacian convolutions
Laplacian edge detector
Laplacian kernels, usage
Laplacian operator
Laplacian weighting function
Large-scale edge information
Laser mapping systems, principle
Laser pulse. *See* Single laser pulse
Laser scan data
Laser scan data, aerial photography (combination)
Laser scanning:
  advantages
  term, usage
Laser scanning systems
  classification
  hardware
  oscillating mirror laser scanning system, scan pattern
  principles
  VMX450 mobile laser mapping system (RIEGL)
Latent image
Latitude:
  forward conversion
  illustrations
  isolation
  values
Laussedat, Aimé
Layers (information)
LCS. *See* Liquid crystal shutter
Leaf-type shutter, schematic diagrams
Least squares. *See* also Moving least squares
  adjustment
    initial approximations
    manual methods
    matrix methods
  application
  definitions
  examples
  fundamental condition
  matching
    iterative process

subarrays, positions
methods
solution
techniques
Left-handed coordinated system
Legs, centerlines (intersection)
Leica:
ADS80 airborne digital sensor
PAV 30
PAV 80 gyro-stabilized aerial-camera mount
RC30 aerial mapping camera
SD3000 analytical plotter
Length equivalents
Lenses. *See also* Cameras
analysis, simplification
aperture diameter, increase
biconvexity
combination lens
cone assembly
converging lens
cost
distortions
correction
falloff
focal length
formula
incident light ray, refraction
light rays, impact
light-gathering power
optical axis
light ray, parallelism
optical glass
plane
plane of infinite focus
resolution (resolving power)
SAGA-F lens, cross-section
simple lenses, usage
speed
stereoscope
thick lens, nodal points
vignetting
Leveling screws (direct optical projection stereoplotters)
Lever arm parameters
facilitation
Light Detection and Ranging (LiDAR)
elevations
Light meter, usage

Light rays:
directions, change
entry
intersection
creation
passage
refraction
Light velocity
Light waves:
frequency/amplitude/wavelength
point source emanation
Line pairs
Line perspective
Linear advance
Linear array imagery
Linear array sensors
image
illustration
scans
Linear drift
Linear features
Linear regression
computation, tabular form
Linear sensory array, orientation fixation
Linear stretch
contrast enhancement, digital image
Linearized collinearity equations, solution
Line-pair test pattern
Lines
Liquid Crystal Display (LCD) computer monitors
Liquid crystal shutter (LCS):
display, stereoviewing principle
glasses, requirement
rate alternation
Local polynomial surface
Local vertical coordinates
conversion
system
Local vertical origin:
translation
X'Y'Z' coordinate system, translation
Longitude:
calculation
forward conversion
values
Loosely coupled integration level
Loss of lock, problem

Low oblique digital photograph, example
Low oblique images
Low oblique photograph, example
Low spatial frequency, density trace (presence)
Low-pass filter
  application
  convolution, result

M

Main frame (direct optical projection stereoplotters)
Manhole, aerial photographs (examples)
Manual of Geographic Information Systems
Manual of Photogrammetry
  camera definition
  K computation
Manual of Photographic Interpretation
Manual of Remote Sensing
Manuscript map, placement
Mapping cameras
  three-dimensional coordinate system
Mapping data collection
Maps. *See also* Flight map
  editing
  manuscripts
  plotting
  projections
  scale
    concept
    optimum, selection
Material costs
Matrix (matrices):
  bandwidth
  equations, analytical self-calibration
  inverses, relationship
  product, formation
  study
Maximum-line-count method
Meades Ranch
Mean sea level (MSL)
Measured image coordinates, refinement
Mechanical projection
  stereoplotters
Mechanical projection stereoplotter
Mechanization. *See* also Inertial navigation system
  involvement
Median filtering
Medium, light rays (passage)

Megapixel
Meter-candle seconds
Metric engineer’s scales
Metric photogrammetry
Metric qualities
Metric system, preference
Metric terrestrial camera
Micrometer, usage
Microsoft UltraCam Eagle ultra-large digital aerial photogrammetric camera
Mirror stereoscope
    operating principles
Mirrors:
    back-surfaced mirror
    first-surface mirror
    plane mirrors, usage
Mistakes
Mixed calculations
Mobile TLS, usage
Model
    base, adjustment
    scale, change
    tilt, correction
Model height ratio, perception
Modified collinearity equations, usage
Modulation transfer:
    determination
    quantification
Modulation transfer function (MTF)
    curve
    Nyquist frequency, relationship
Monocomparators, usage
Monoscopic depth perception
Mosaics
    construction
    controlled digital mosaics
    semicontrolled digital mosaics
    semicontrolled mosaic, example
    uncontrolled digital mosaics
    work
Most probable value (MPV)
    definition
Mount
Moving least squares
Moving windows
    convolutions
    edge detection
    image convolution

operations
inputs
3×3 position, convolution
usage
MPV. *See* Most probable value
MSL. *See* Mean sea level
MTF. *See* Modulation transfer function
Multicollimators
Multiple reflections, creation
Multiplication, commutative property
Multipurpose GIS, usage
Multipurpose land information system, photogrammetric applications
Multiscale representation
Multispectral imagery

N

NAD. *See* North American Datum
NADCON
Nadir:
datum nadir point
ground nadir point
location (oblique photograph)
photographic nadir point
point, fiducial coordinates (computation)
vertical lines, intersection
Nadir linear array sensor scans
NAPP. *See* National Aerial Photography Program
NASA. *See* National Aeronautics and Space Administration
National Aerial Photography Program (NAPP)
National Aeronautics and Space Administration (NASA)
National Flood Insurance Program (NFIP), FEMA management
National Geodetic Vertical Datum of 1929 (NGVD29)
National High Altitude Photography (NHAP)
aerial photos
Program
National map accuracy standards (NMAS)
National Map Accuracy Standards, meeting
National Spatial Reference System (NSRS)
National wetlands inventory (NWI)
NAVD88. *See* North American Vertical Datum of 1988
Nearest neighbor
interpolation, shape
method, discussion
Near-vertical aerial photographs, images
Near-vertical photographs
affine transformation, usage
stereopair

Near-vertical photography, assumptions
Near-visible spectrum, white light (dissection)
Neat models, transparent templates (usage)
Negative
    ownership
    production
    reversal
Negative plane
NGVD29. *See* National Geodetic Vertical Datum of 1929
NHAP. See National High Altitude Photography
Niepce, Joseph
NMAS. See National map accuracy standards
Nodal points
Noise filter method
Nominal, term (usage)
Nonbanded reduced normal equations
Nonengineering applications, elements
Nonimage data
Nonimage data, image data (combination)
Nonlinear collinearity equations, linearization
Nonmetric cameras, calibration
    complication
Nonmetric terrestrial camera
Nonorthogonality, correction
Nonreferenced digital image, schematic representation
Nonreferenced satellite image
Nonzero coefficients, inclusion
Nonzero digits, significance
Nonzero submatrices
    appearance, cause
Normal color emulsions, usage
Normal equations
    formulation
Normalized cross-correlation
    coefficient
    correlation coefficient
North American Datum of 1927 (NAD27)
North American Datum of 1983 (NAD83)
    readjustment
North American Vertical Datum of 1988 (NAVD88)
NSRS. *See* National Spatial Reference System
Numerical relative orientation, calculation
NWI. *See* National wetlands inventory
Nyquist frequency
    modulation transfer function, relationship
Nyquist rate

O

Objects:
- apparent depth
- close-range stereo coverage
- density distributions
- density modulation
- depth, importance
- distance
  - variation
- heights, parallax differences (equation)
- information
- inverted image
- location/marking, importance
- points, photo coordinates
- position, displacement
- relative sizes
- vertical heights

Object-space coordinates
- conversion sequence
- standard deviations, indications
- systems
  - conversions
  - image coordinate system, parallelism

Object-space positions, establishment
Oblique, nonrectified photograph, digital image
Oblique aerial photographs, exposure
Oblique photographs
- coordinate axis system
- horizon, location
- horizontal angles, computation
- nadir, location
- vertical angles, computation

Oblique photography, camera axis (inclination angle determination)
Oblique terrestrial photograph:
- exposure
- measurements, horizontal/vertical angles

Observations:
- definition
- equations

Off-diagonal nonzero element
Omega
- angle, x’ axis (rotation)
- direct optical projection stereoplotters
- rotation

Omega-phi-kappa *(ω-ϕ-κ)*:
- angular orientation
- equations, usage

kappa, rotation
phi, rotation
system, photograph (orientation)
One-point perspective image
On-screen digitizing
On-the-fly (OTF) ambiguity resolution, development
On-the-fly (OTF) integer ambiguities
On-the-fly (OTF) techniques
Operator controls, interface
Optech Orion airborne laser scanning system
Optical axis (lens)
Optical-mechanical projection
Optical-mechanical rectification
Optics
Orientation. *See* also Absolute orientation; Angular orientation; Interior orientation; Relative orientation
analytical plotters
completion
fixes
establishment
procedures, softcopy plotters
software
Origin, vector (relationship)
Orthogonal matrix
Orthophotomaps
Orthophotos
digital orthophoto production process, schematic representation
generation
production
usage
Orthorectification:
DEM accuracy, effect
differential rectification
Oscillating mirror:
comparison
laser scanning systems, scan pattern
OTF. *See* On-the-fly
Outward radial distortion
Overbar, usage
Overlapping aerial photography, geometry
Overlapping pair, geometry (vertical photographs)
Overlapping vertical photographs
Over-parameterization, occurrence

P

P-3 analytical plotter (Zeiss)
Panchromatic emulsions

Panel points
requirements
Paper media, stability
Parabolic antenna, shape (determination)
Parallactic angle
changes, detection
increase, B/H ratios (increase)
perception
stereoscopic depth perception
Parallax. *See* also Stereoscopic parallax; X parallax; Y parallax
differences
usage
equations
monoscopic measurement
variation
wedge
Parallax bar
constant, determination
schematic diagram
Parallax differences:
elevations
example
measurement
objects, heights (equation)
Parallax measurement
flight axes, usage
monoscopic methods
photographic flight-line axes
simple scale, usage
stereopairs, orientation
stereoscopic methods
advantages
Parallelepiped:
formation, vectors (usage)
result
Parking lot images:
convolution, Laplacian kernels (usage)
high-pass filter, application
image pyramid, formation
large-scale edge information
low-pass filter, application
median filter, application
noise
Sobel edge detection, application
Partial derivatives:
coefficients, letter substitutions
nonzero values

Pass points:
configuration
generation
locations
measurement
PAV 30 (Leica)
PAV 80 gyro-stabilized aerial-camera mount (Leica)
PE. *See* Percent end
Pencil of rays
Percent end (PE) lap
Percent side lap, adjustment
Percentage of end lap (PE)
Perspective geometry, laws
Perspective projection
provision
Perspective-projection transformation
matrix
Phi
angle, y_1 axis (rotation)
direct optical projection stereoplotters
rotation
Photo. *See* Photographs; Photography
Photo control:
establishment
ground-surveyed network
images
number/location
object-space positions, establishment
Photo coordinates. *See* Photographic coordinates
Photogrammetric control:
artificial targets, usage
establishment, field surveys (usage)
field surveying, two-step procedure
points
Photogrammetric Engineering and Remote Sensing (ASPRS)
Photogrammetric flight plans
Photogrammetric ground control survey, planning (importance)
Photogrammetric mapping, accuracy capabilities
Photogrammetric operations, bridging (usage)
Photogrammetric problems, analysis/solution
Photogrammetric products, scales/accuracies
Photogrammetric Record
Photogrammetric Record
Photogrammetric scanners
Photogrammetry:
definition
geographic information systems, relationship

history
imaging device, usage
interpretative photogrammetry
metric photogrammetry
professional photogrammetry organizations
techniques, advancements
tool
usage
uses
Photographic base-height ratio (B/H)
Photographic coordinates
comparator measurement
correction
measurements
accuracy
analytical methods
constant terms
scales, usage
pairs, light ray representation
refinement
system, usage
weight matrix
Photographic distance, histogram
Photographic emulsions:
black-and-white emulsions, sensitivities
characteristic curve
characteristics
color emulsions, layers
darkness, degree
density
exposure
grains
infrared (IR) emulsions
panchromatic emulsions
spectral sensitivity
Photographic end lap
Photographic film:
color film
color infrared film (false-color film), sensitivity
components
cross section
distortions
film-advancing mechanisms
film-based aerial mapping cameras
film-flattening mechanisms
flattening
mechanisms, challenge

photographic film, cross section
sensitivity/speed (indication), ISO number (usage)
speed. See Aerial film speed.
total exposure
Photographic flight-line axes
Photographic measurements glass scales
scales
Photographic nadir point
Photographic papers, distortions
Photographic photogrammetry
Photographic side lap
Photographs
adjustment
camera calibration parameters, loosening (examples)
example
input parameters
output
t-statistics
angular attitude, INS capability
black-and-white photographs, processing/printing
block
configuration
graphical representation
classifications
connections
digital coverage
distance
intersection
end lap
flight lines, location/marking
high oblique photograph, example
low oblique digital photograph, example
low oblique photograph, example
obscured area
orientation
problem
stereoscopic orientation
photo base
determination
stereopair
pseudoscopic view
rectification
scale
shared pass points, connections
shrinkage/expansion
spatial position/angular orientation
stereopair, epipolar geometry

stereoscopic viewing
strips, control configuration (example)
superposition
tilt, problem
two-photo adjustment, convergence
types
vertical aerial photographs, taking
Photography:
index
pictorial quality
purpose
scale
Photomaps
advantages
disadvantages
preparation
Phototheodolites
Phototriangulation
Physical optics
Pictorial qualities
considerations
Picture elements (pixels)
color, representation
input
megapixels
pattern
superimposition
Picture format, diagonal
Pinhole box
Pinhole camera, principle
Plane, perspective projection
Plane mirrors, usage
Plane of infinite focus
Planimetric accuracy
Planimetric correctness, impact
Planimetric details, location
Planimetric feature:
digitization
direct compilation, stereoplotter (usage)
tracing
Planimetric mapping, reflection instruments (usage)
Planimetric maps
Planimetric maps, compilation
Planimetry:
inconsistencies
renditions
Platen (direct optical projection stereoplotters)

vertical movement
Plotters:
analytical plotters
softcopy plotters
stereoplotters
Pocket stereoscope
cost
photographs, orientation
schematic diagram
usage
Point clouds
Point names, subscripts (correspondence)
Point of best symmetry
Point perspective
Point positioning
errors
Point source:
bundle rays, emanation
light wave emanation
Points
coordinates, transformation
Point-spread function
Polar stereographic Mercator
Polarized-platen viewing (PPV) system
anaglyphic system, comparison
Polarizing filter screen:
display, stereoviewing principle
imaging, softcopy photogrammetric workstation
Polyconic Mercator
Polygons
Polynomial error curves
Polynomial transformation, usage. *See* Three-dimensional polynomial transformation
Positions:
measurement
examples
reset, INS mechanization (usage)
Positive print, obtaining
Post-bundle adjustment statistics
PPV. *See* Polarized-platen viewing
Precise (P) code
Precise emphemeris
Precision:
concepts
matrix equations
Predicted ephemeris
Preprocessing operations
Preprocessing usage

Principal distance (direct optical projection stereoplotters)
Principal line
Principal plane
    side view
        tilting-lens optical-mechanical rectifier, schematic diagram
Principal point
    conjugate principal points
    coordinates
        relationship
    definitions, images
    distance
    example
    location
    origin, coordinates (reduction)
    parallax
    photo coordinates
    transfer, floating mark principle (usage)
Processed imagery (edges), sensor tilt (correspondence)
Processed linear array imagery
Professional photogrammetry organizations
Project area, example
Project complexities, variations
Project planning
    cost estimation/scheduling
    photography, purpose
    specifications
Projection distance (direct optical projection stereoplotters)
Projection printer, enlargement (geometry)
Projection printing
Projection systems, direct optical projection stereoplotters
Projective coordinate transformation
Projective transformation, complication
Projectors:
    bar tilt
    direct optical projection stereoplotters
    motions
Project-specific basis, list
Property management, photogrammetric applications
Pseudoranges
Pseudoscopic view
Pulfrich, Carl
Pulse spacing
Pupil
Pushbroom scanners
Pythagorean theorem

Quantitative photogrammetric measurements, requirement
Quantization
QuickBird, IKONOS (contrast)

R

Radial distance:
    change
    illustration
Radial distortions
Radial lens distortion/corrections, relationship
Radiometric resolution
Radiometric scale
Random errors
    normal distribution
Random Sample Consensus (RANSAC)
Rapid static method
Rational polynomial coefficient (RPC)
    camera models
    model, sufficiency
Raw image
Raw linear array imagery
Rays, bundle
    emanation
RC30 aerial mapping camera (Leica)
RCI. *See* Roadway Characteristics Inventory
Rear nodal point
Receivers, components
Rectangular coordinates (measurement), algebraic signs (usage)
Rectangular DEM, three-dimensional representation
Rectangular project areas, coverage
Rectangular-to-polar conversion capability
Rectification
    analytical rectification
    control points, plot
    differential rectification (orthorectification)
    digital rectification
    geometry, display
    ground control points, relief (correction)
    optical-mechanical rectification
    tilted photographs
    two-dimensional projective transformation, usage
Rectified photographs
Rectified-ratioed coordinates
Rectified-ratioed photo, points (projection)
Rectifiers
Rectifying latitude
Red Green Blue (RGB) color space

Reference ellipsoids:
definition
parameters
Reference mark
Reference photo coordinate system, establishment
Reference table (direct optical projection stereoplotters)
Reference variance
Reflection instruments, usage
Refraction, angle
Refractive index
Regular grid, three-dimensional view (interpolation)
Relative frequencies
Relative orientation
analytical plotters
direct optical projection stereoplotters
numerical relative orientation, calculation
points, usage
Relative positioning
carrier-phase methods
Relief displacement
corrections, rectification (control point plot)
evaluation, equation
impact
reduction
tilted photograph
vertical photograph
Relor program, running
Remote sensing:
branches
image classification, performing
Resampling:
bilinear/bicubic interpolation
nearest-neighbor method
Resect program, running
Residual:
definition
squares, sum (minimization)
Resolution (resolving power). *See also* Geometric resolution; Radiometric resolution; Spectral resolution
calibration
importance
test pattern
Retina
Retina, sensitivity
Reverse transformation, obtaining (problem)
RGB. *See* Red Green Blue
RIEGL VMX450 mobile laser mapping system

RIEGL VZ400 terrestrial laser scanner
Right-handed coordinated system
RMK TOP 15 aerial mapping camera (Zeiss)
RMS. *See* Root mean square
Roadway Characteristics Inventory (RCI)
    GIS, example
Roll diversity
Root mean square (RMS):
    calculation
    error
Rotated x′y′z′ image coordinate system, measurement
Rotating-disk lens shutters
Rotation:
    angles, computation
    azimuth/tilt/swing terms
    equations, development
    omega/phi/kappa terms
    parallel relationships
    parameters, obtaining
    systems, conversion
    X″ axis
    Z′ axis
Rotation matrix (matrices)
    combination
    computation
    formation
    multiplication
Rover
    receiver, movement
Row coordinate, usage
Row interpolation
    weight matrix, computation
RPC. *See* Rational polynomial coefficient

S

Safe lights, usage
SAGA-F (usage)
SAGA-F lens, cross-section
Sample correlation coefficient
Sampling frequencies, effect
Sanitary sewer facility map, example
Satellite image triangulation
Satellite lock, loss
Satellite photos, usage
Saturation:
    polar coordinates
    representation

color value basis
Scale
application
change
relationship
Scale factor
quantification
Scaled X′Y′ coordinate system, superimposition
Scan angle
Scanner, geometric quality
Scanning-type satellite image, two-dimensional affine relationship
Scheimpflug condition
SD3000 analytical plotter (Leica)
Search array, example
Season, impact
Section lines
Sediment yield, GIS analysis
Self-calibrating bundle adjustment
Self-calibration adjustment, solution approach
Semianalytical aerotriangulation
advantage
control
fundamentals
involvement
pass points, configuration
Semicontrolled aerial mosaics
Semicontrolled digital mosaics
control/tie point configuration
Semicontrolled mosaic, example
Sequential angular rotations
Sequential independent model triangulation, usage
Sequential strip formation
Servomotors:
impact
usage
Seven-model strip, control extension (plan view)
Shadows
digital image
Shared pass points, connections
Ship propeller size/shape, determination
Shoulder region
Shutter speed:
aperture, relationship
f-stop, relationship
variation
Shutters
efficiency

leaf-type shutter, schematic diagrams
light regulation
priority mode
SIA. *See* Stereo-image alternator
Side fiducial:
mark
photographic coordinate system
Side lap. *See* also Photographic side lap
definition
flight strip photograph
overlap
Signal noise
Significant figures
determination
number, equivalence
value
Signs/lights, project-specific basis
Silver halide crystals
exposure
grains
Silver halide layers
Simple lenses, usage
Simultaneous bundle adjustment
Simultaneous Multi-camera Analytical Calibration (SMAC)
Sinc function, form
Single laser pulse, multiple returns
Single-lens camera
image formation
light pencils
Single-lens frame cameras
Single-lens reflex camera, example
Single-receiver point positioning
68 percent error
SMAC. *See* Simultaneous Multi-camera Analytical Calibration
Snell's law
SNWA. *See* Southern Nevada Water Authority
Sobel edge detection, usage
Softcopy instruments, innovation
Softcopy mapping systems, image classification processes (usage)
Softcopy photogrammetric workstation, polarizing filter stereo imaging
Softcopy plotters
absolute orientation
advantage
coplanarity
epipolar geometry
epipolar plane
hinge point

image measurements
LCD, physical construction
manual usage
orientation procedures
system hardware
template
Softcopy stereoplotters
impact
Solid-state CCD imaging array
Southern Nevada Water Authority (SNWA)
Space intersection
collinearity, usage
stereopair (aerial photos)
Space oblique Mercator
Space resection
collinearity, usage
Spatial data collection
Spatial domain
Spatial frequency
digital image
variation
Spatial image, abstraction
Spatial position, expression
SPC. *See* State plane coordinate
Special-purpose topographic maps
Specific force, impact
Spectral resolution
Spectral transformations
Spectrum:
electromagnetic spectrum, wavelength classification
near-visible spectrum, white light (dissection)
visible spectrum, white light (dissection)
Spider, usage
SPOT. *See* Système Pour d'Observation de la Terre
Spot elevations, weighted average
Square camera focal-plane format
Square grid, tilted photo image
ST-4 mirror stereoscope
binocular attachment/parallax bar
Standard deviation, definition
Standard parallels
State plane coordinate (SPC) systems
Static terrestrial photography
Statistics, involvement
Statue of Liberty, digital image
quantization levels, variation
spatial resolutions, variation

Stereo display configuration
Stereo workstation, computer service
Stereocomparators, usage
Stereo-image alternator (SIA)
    synchronized shutters, usage
Stereometer
Stereometric cameras
Stereomodels
    analytical stereomodel
    arbitrary coordinate systems
    continuous strip
    level, problem
    manual interior orientation
    orientation, control (recommendation)
    pass points, locations
    perception
    planimetric features, digitization
    relative orientation
    scaling, horizontal control points (requirement)
Stereopairs
    analytical relative orientation
        performing
    automatically matched DEM, positions
    epipolar geometry
    images, usage
    obtaining
    orientation
    photo base
    space intersection
    taking
Stereoplotters (plotters)
    characteristics
    evolution
    mechanical/optical limitations
    projectors, comparison
    usage
Stereoscopes
    ST-4 mirror stereoscope
    usage
    Zoom 95 stereoscope
Stereoscopic coverage failure
    flying height variations, impact
    terrain variations, impact
    tilt, impact
Stereoscopic depth perception
    parallactic angle
Stereoscopic impression, creation

Stereoscopic methods
Stereoscopic model
Stereoscopic neat model, area coverage
Stereoscopic overlap area
Stereoscopic parallax
  error evaluation
  measurement, flight-line axes (usage)
  problems
  vertical aerial photographs
Stereoscopic photography, effectiveness
Stereoscopic plotters, classification
Stereoscopic plotting instruments
  design concept
  stereomodels (orientation), control (recommendation)
Stereoscopic viewing
  example
  facilitation
  geometry
  photographs, orientation
Stereoviewing
  base-height ratio (b_e/h)
    estimation
  display monitor, usage
  principle (alternating LCS shutter display)
  principle (polarizing filter screen display)
Stop bath, black-and-white photographic processing/printing step
Straight-line portion
Streets/highways, project-specific basis
Strip:
  adjustment
  cameras
  chain method, usage
  control extension, XYZ coordinates (error accumulation)
  cross-strip ordering
  down-strip ordering
  error propagation
  full strip, rotations
  lateral advance
  photo block, configuration
  photo number
  seven-model strip, control extension (plan view)
Strip model:
  adjustment
  sequential construction, independent models (usage)
Structures, project-specific basis
Subarrays
  average digital number, computation

digital numbers
pixels
positions
size, results
Sub-classes
Sub-images, usage
Support, photographic film component
Surface splines
Surveying and Land Information Science
Swaths
Swing:
angular orientation
definition
rotation
angles
example
Symmetric radial distortion (decentering)
computation
pattern
Symmetric radial lens distortion
coefficients
Systematic errors
elimination
measurement
Système Pour d’Observation de la Terre (SPOT)
satellite, usage
sensor systems

T

Tangential lens distortion
Targets. *See* also Artificial targets
sizes, design
Taylor, Brook
Taylor’s series
application
first-order terms
usage
Taylor’s theorem, usage
Telecommunication tower sites, example
Template (softcopy plotters)
Terrain:
areas
elevation
representation, contour map (usage)
three-dimensional pictorial view, DEM/orthophoto (usage)
variations, impact
Terrestrial cameras

types
Terrestrial laser scanning (TLS)
data, georeferencing
VZ400 terrestrial laser scanner (RIEGL)
Terrestrial photogrammetry
advantage
applications
control
prominence, continuation
usefulness
Terrestrial photographs
Terrestrial photography
control
types
Test object, (low spatial frequency), density trace (presence)
Thick lens, nodal points
3dconf program, usage
Three-dimensional conformal coordinate transformation
achievement
approximations, requirement
initial approximations
swing calculation
tilt/azimuth, calculation
tilt/azimuth/swing, combination
vector computation
points, rotation
rotation
scaling
sequential independent model triangulation, usage
solution, complexity
translation
usage
Three-dimensional control points, requirement
Three-dimensional coordinates
system
Three-dimensional digitizer
Three-dimensional ellipsoid, formation
Three-dimensional ground coordinates
cubic polynomial function ratio, usage
transformation
Three-dimensional impression, creation
Three-dimensional object space coordinates, determination
Three-dimensional operation:
analytical plotters
analytical stereoplotters, flowchart
Three-dimensional polynomial transformation, usage
Three-line linear array sensor scans

Three-line linear sensor array, orientation fixation
Three-line scanner:
    aerotriangulation
    geometry
    image scene collection
Three-point perspective image
Three-times-rotated coordinate system
Tie points:
    approximations, discovery
    configuration
    coordinates, photographs (connections)
    identification
    usage
Tightly coupled integration level
Tilt
    angular orientation
    calculation
    impact
    orientation angle
    rotation
        angles
        example
Tilt angle:
    determination
    display
Tilted aerial photographs, atmospheric refraction
Tilted photographs
    auxiliary tilted photo coordinate system
    auxiliary x′y′ image coordinate system
    effect
    geometry
    image displacements
    image point
    principal plane
        rectification, geometry (display)
    rectification
    relief displacement
    scale
    example
    square grid, image
    xy fiducial system, x′y′-tilted system (conversion)
Tilting-lens optical-mechanical rectifier, schematic diagram
Tilt-swing-azimuth (t-s-*a*):
    angular orientation, display
    system
TIN. *See* Triangulated irregular network
TLS. *See* Terrestrial laser scanning

Toe region
Topographic features, representation
Topographic mapping
Total exposure
    diaphragm area, shutter speed (product)
Tracing pencil (direct optical projection stereoplotters)
Tracing systems, direct optical projection stereoplotters
Tracing table (direct optical projection stereoplotters)
Traffic accident investigations, photogrammetry (usage)
Traffic management, photogrammetry (usage)
Transformation:
    affine transformation, usage
    coordinate transformation approach
    homogeneous coordinates, usage
    parameters
        coefficients, matrix
        computation
        usage
    performing
    projective transformation, complication
Translation factors:
    addition
    calculation
Translation parameters
Transparencies (diapositives)
Transportation:
    photogrammetric applications
    project-specific basis
Transverse Mercator
    coordinate zone, defining constants
    coordinates, conversion
    map projection
    projections, setup
Traversing
Triangle, h (altitude)
Triangulated irregular network (TIN)
    data points
    generalization
    models, construction
Triangulation. *See* also Delaunay triangulation
    horizontal angles, measurement
Trigonometric functions, angles (involvement)
Trigonometric leveling
    usage
Tri-lap area
    points, measurement
Trilateration

Triple-scalar product, zero value
True value, definition
Truly vertical overlapping photographs
Truly vertical photographs
T-statistics, examples
Two-dimensional, term (usage)
Two-dimensional affine coordinate transformation
    example
    nonorthogonality, correction
    relationship
    rotation
    translation
    x, scale change
    y, scale change
Two-dimensional affine relationship
Two-dimensional affine transformation
Two-dimensional conformal coordinate transformation
    method, alternative
    performing
    rotation
    scale change
    steps
    translation
Two-dimensional conformal transformation
Two-dimensional coordinate transformation parameters
Two-dimensional ellipse
Two-dimensional homogeneous points, embedding
Two-dimensional projective coordinate transformation
Two-dimensional projective rotation matrix, parameterization
Two-dimensional projective transformation:
    equations, form
    geometry
    representation
    rotation, parallel relationships
    usage
Two-dimensional rational polynomial coordinate transformation
Two-photo adjustment, convergence
Two-point perspective image

U

UltraCam Eagle ultra-large digital aerial photogrammetric camera (Microsoft)
Uncontrolled aerial mosaics
Uncontrolled digital mosaics
    construction
Unexposed emulsion, fog
Unfiltered ALS data
Unit equivalents/fractions

Unit weight, standard deviation
Units
Universal Transverse Mercator (UTM):
    projection
    system
Urban scenes, two-dimensional transforms
U.S. Army of Corps of Engineers
U.S. Department of Agriculture, aerial photography resource
U.S. Geological Survey (USGS):
    aerial photograph archive
    camera calibration
    parameters
    quadrangle maps, usage
    satellite image archive
U.S. survey foot
Utilities, project-specific basis
UTM. *See* Universal Transverse Mercator

V

Vanishing points
Variable terrain:
    ALS system, swath pattern
    vertical aerial photograph, usage
    vertical photograph, scale
Variances
    weight, inverse proportion
Vector
    azimuth
    orthogonal
    superimposition
Vegetation layer:
    example
    usage
Velocity
    reset, INS mechanization (usage)
VERTCON
Vertical accuracy
Vertical aerial photographs:
    example
    scale, example
    stereoscopic parallax
    taking
    usage
Vertical angles, computation (oblique photographs)
Vertical control
    establishment, field survey methods (usage)
    images

point, usage
points
surveys
differential leveling, usage
Vertical datums
Vertical exaggeration:
analysis, diagrams
calculation
magnitude, expression
stereoviewing
Vertical ground-surveyed photo control points
Vertical photographs
example
flight line
flying height
geometry
ground coordinates
example
overlap
overlapping pair
geometry
relief displacement
equation
example
photographic example
scale
determination
variations
stereopair, air base
two-dimensional view
Vertical plane, horizontal angle
Viewing systems, direct optical projection stereoplotters
Vignetting
Visible light, wavelengths (discernment)
Visible spectrum, white light (dissection)
VMX450 mobile laser mapping system (RIEGL)
VR mapping workstation
VZ400 terrestrial laser scanner (RIEGL)

W

Washing, black-and-white photographic processing/printing step
Water quality management:
GIS project, location (example)
photogrammetric applications
Wavelength
Wavelet decomposition. *See* First-level wavelet decomposition
Wavelet transform

Weather conditions
Weather conditions, problems
Weight, definition
Weighted observations
    equations
    system
Wetland analysis, photogrammetric applications
Wetland area, GIS database (example)
WGS. *See* World Geodetic System
WGS84 datum, usage
Whiskbroom scanners
White light, dissection
Wide-angle lenses, illuminance (decrease)
Wildlife management:
    GIS project
        study area
        vegetation layer
    photogrammetric applications
World Geodetic System of 1984 (WGS84)
Wrap-around effect, creation

X

x’ axis, omega rotation
X″ axis, rotation
    example
X coordinates, usage
X ground coordinates
X parallax
X plate coordinates, measurement
X value
XY coordinates
    forward/direct conversion
xy ground coordinate system, transformation
XY two-dimensional coordinate system
x′y′-tilted system
X′Y′Z′ coordinate system, translation
XYZ ground coordinates, computation
xyz image coordinate systems, measurement
XYZ model coordinates, computation
XYZ right-handed three-dimensional coordinate systems
xyz right-handed three-dimensional coordinate systems
XYZ stereoplotter controls

Y

Y coordinates, usage
Y ground coordinates
    standard deviations

Y parallax:
causes
flying height, variation (impact)
photographs
orientation, problem
tilt, problem
Y plate coordinates, measurement
Y value
y_1 axis, phi rotation

Z

Z′ axis, rotation
example
Z coordinate:
determination
standard deviation
Z_2 axis, kappa rotation
Zeiss P-3 analytical plotter
Zeiss RMK TOP 15 aerial mapping camera
Zenith, vertical lines (intersection)
Zeros, significance
Zigzagging pattern, generation
Zoom 95 stereoscope
Zoom stereoscope
Zoom transfer scope (ZTS)
usage

image

FIGURE 2-21 (*a*) Normal color image and (*b*) color infrared image. Note that healthy vegetation, which appears green in the normal color image, appears red in the color infrared image. Circled tennis courts are painted green but appear gray in the color infrared image.

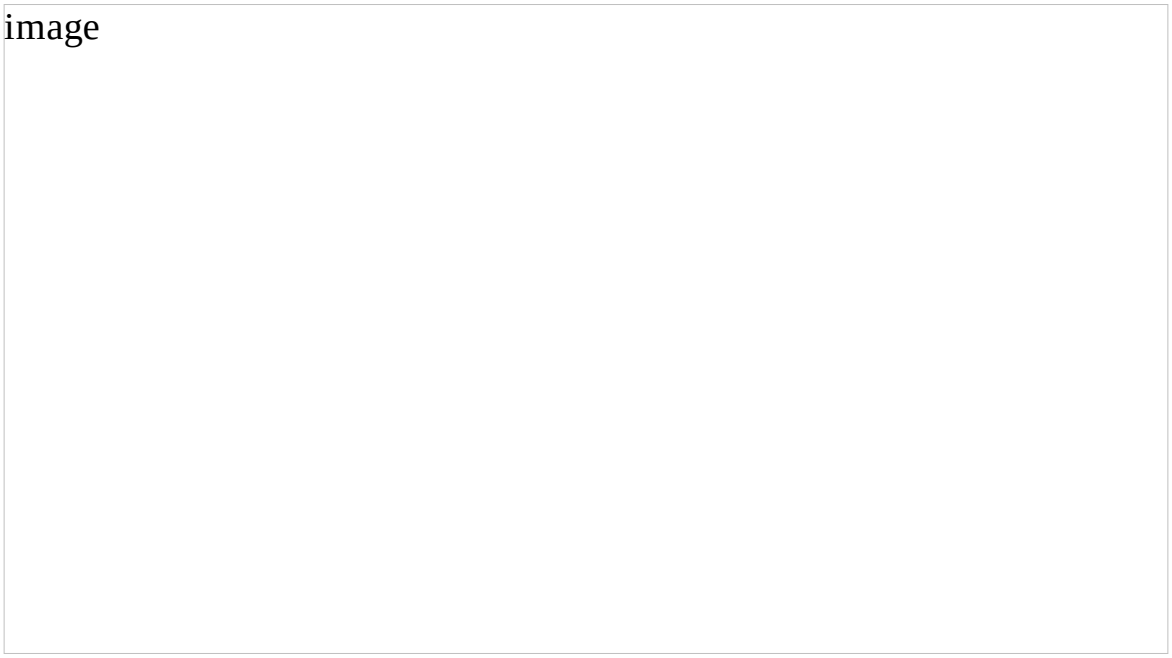


FIGURE 2-27 (*a*) A view of the color cube from behind the origin and (*b*) a view of the color cube from the opposite corner.

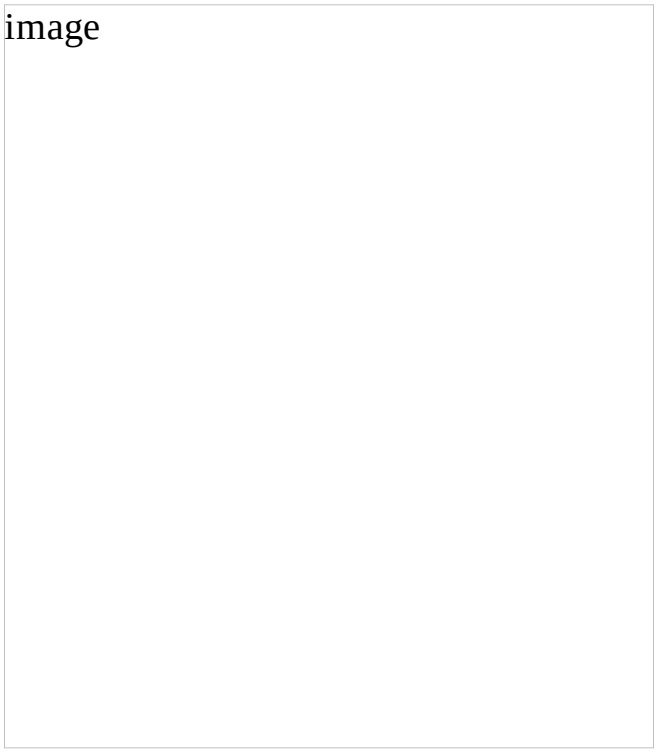


FIGURE 2-30 Representation of hue and saturation corresponding to Fig. 2-29.

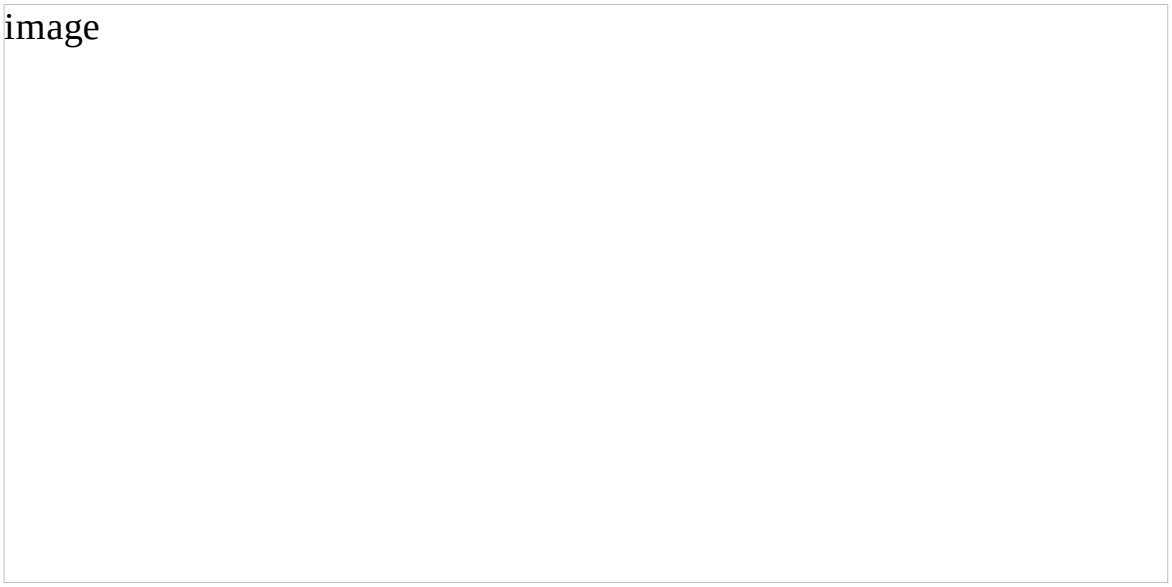


FIGURE 2-31 Illustration of the (*a*) color additive process and (*b*) color subtractive process.

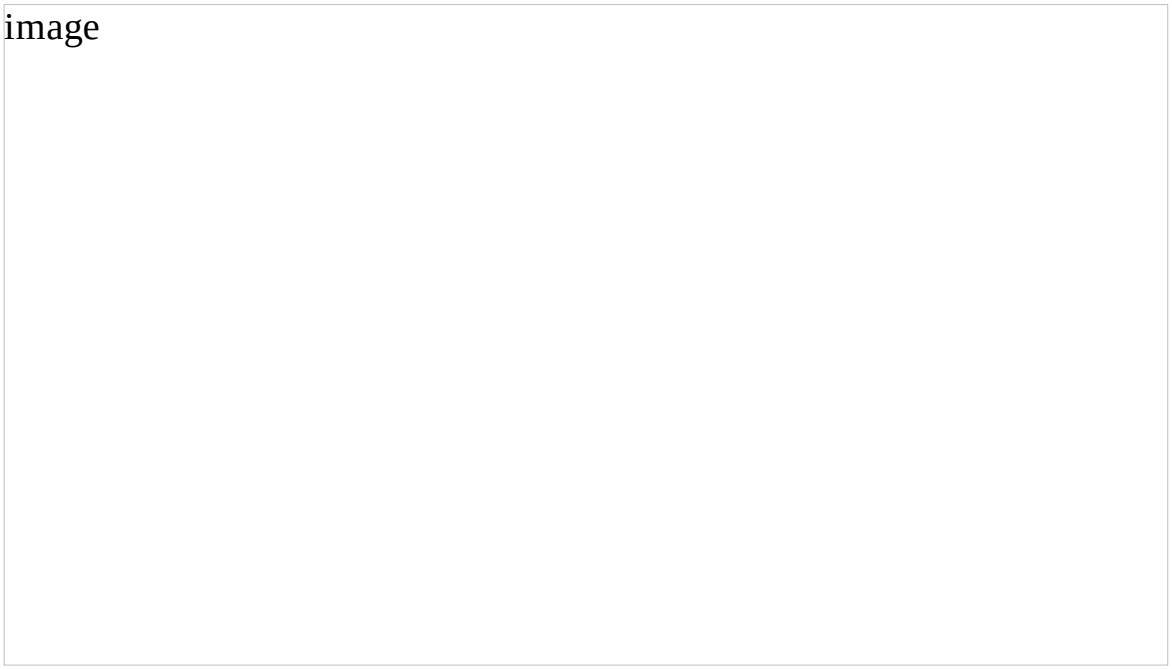


FIGURE 3-9 Solid-state CCD imaging array of 14,600 × 17,200 (250 million) pixels. (*Courtesy Teledyne DALSA.*)