56

Photogrammetry and Remote Sensing

J.S. Bethel
Purdue University

56.1 Basic Concepts in Photogrammetry

The term *photogrammetry* refers to the measurement of photographs and images for the purpose of making inferences about the size, shape, and spatial attributes of the objects appearing in the images. The term *remote sensing* refers to the analysis of photographs and images for the purpose of extracting the best interpretation of the image content. Thus the two terms are by no means mutually exclusive and each one includes some aspects of the other. However, the usual connotations designate geometric inferences as photogrammetry and radiometric inferences as remote sensing. Classically, both photogrammetry and remote sensing relied on photographs, that is, silver halide emulsion products, as the imaging medium. In recent years digital images or computer resident images have taken on an increasingly important role in both photogrammetry and remote sensing. Thus many of the statements to be found herein will refer to the general term *images* rather than to the more restrictive term *photographs*. The

characteristic of photogrammetry and remote sensing that distinguishes them from casual photography is the insistence on a thorough understanding of the sensor geometry and its radiometric response.

The predominant type of imaging used for civil engineering applications is the traditional 23-cm format frame aerial photograph. Photogrammetric and remote sensing techniques can be equally applied to satellite images and to close-range images acquired from small format cameras. The main contributions of photogrammetry and remote sensing to civil engineering include topographic mapping, orthophoto production, planning, environmental monitoring, database development for geographic information systems (GIS), resource inventory and monitoring, and deformation analysis.

Scale and Coverage

A typical aerial camera geometry is shown in Fig. 56.1. The perspective geometry of an ideal camera would dictate that each object point, *A*, would lie along a line containing the corresponding image point, *a*, and the perspective center, *L*. *H* represents the flying height above the terrain; *f* is the focal length or principal distance of the camera or sensor. In the ideal case that the camera axis is strictly vertical and the terrain is a horizontal plane surface, the *scale* of the image can be expressed as a fraction:

$$\text{Scale} = \frac{\text{Image distance}}{\text{Object distance}} \qquad (56.1)$$

FIGURE 56.1 Typical aerial camera geometry.

The units of measure should be the same for the numerator and the denominator so the fraction becomes a unitless ratio. Usually one forces the numerator to be 1, and this is often called the representative fraction or scale. In practice, units are sometimes introduced, but this should be discouraged. For instance, "one inch equals one hundred feet" should be converted as

$$\frac{1 \text{ inch}}{100 \text{ feet}} = \frac{1 \text{ inch}}{1200 \text{ inches}} = \frac{1}{1200} \qquad (56.2)$$

This scale could also be represented in the form 1:1200. By the geometry shown in Fig. 56.1, the scale may also be expressed as a ratio of focal length and flying height above the terrain:

$$\text{Scale} = \frac{f}{H} \qquad (56.3)$$

In practice, images never fulfill the ideal conditions stated above. In the general case of tilted photographs and terrain which is not horizontal, such a simple scale determination is only approximate. In fact the scale may be different at every point, and further, the scale at each point may be a function of direction. Because of this one often speaks of a *nominal scale*, which may apply to a single image or to a strip or block of images.

From Eq. (56.3) it is clear that high altitude imagery, having large *H*, yields a small scale ratio compared with lower altitude imagery, for which *H* is smaller. Thus one speaks of small scale imagery from high altitude and large scale imagery from low altitude. The area covered by a fixed size image will of course be inversely related to scale. Small scale images would cover a larger area than large scale images.

Relief and Tilt Displacement

If the ideal conditions of horizontal terrain and vertical imagery were fulfilled, a scale could be determined for the image and it could be used as a map. Terrain relief and image tilt are always present to some extent, however, and this prevents the use of raw images or photographs as maps. It is common practice to
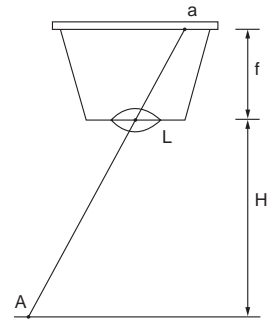
modify images to remove tilt displacement (this is referred to as *rectification*); such images can be further modified to remove relief displacement (this is referred to as *differential rectification*). Such a differentially rectified image is referred to as an *orthophoto*.

The concept of relief displacement is shown schematically in Fig. 56.2. If the flagpole were viewed from infinity its image would be a point and there would be no relief displacement. If it is viewed from a finite altitude its image will appear to "lay back" on the adjacent terrain. The displacement vector in the image is aa'. The magnitude of this displacement vector will depend on the height of the object, the flying height, and its location in the image. Such displacements in the image are always radial from the image of the *nadir point*. The nadir point is the point exactly beneath the perspective center. In a vertical photograph this would also coincide with the image *principal point,* or the foot of the perpendicular from the perspective center. Assuming vertical imagery, the height of an object can be determined by relief displacement to be
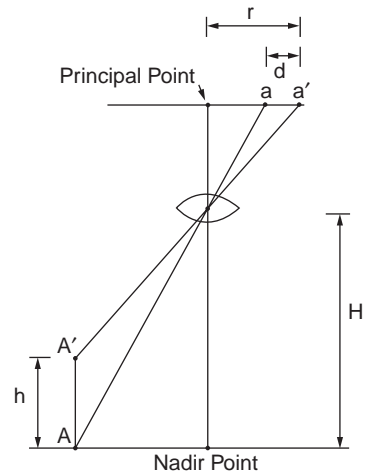

FIGURE 56.2 Relief displacement.

$$h = \frac{dH}{r} \tag{56.4}$$

d and r should be measured in the same units in the image, and h and H should be in the same object space units. Equation (56.4) could be used to obtain approximate heights for discrete objects with respect to the surrounding terrain. The same displacement occurs with the terrain itself, if it is higher than the reference plane. However, because the terrain is a continuous surface this displacement is not obvious as in the case of the flagpole.

Image tilt also creates image point displacements that would not be present in a vertical image. Extreme image tilts are sometimes introduced on purpose for *oblique photography*. For nominally vertical images, tilts are usually kept smaller than three degrees. Figure 56.3 illustrates some of these concepts in a sequence of sketches. Figure 56.3(a) depicts a tilted image of a planimetrically orthogonal grid draped over a terrain surface, Fig. 56.3(b) shows the image with tilt effects removed, and Fig. 56.3(c) shows the image with relief displacement effects removed. Only after the steps to produce image (c) can one use the image as a map. Prior to this there are severe systematic image displacements.
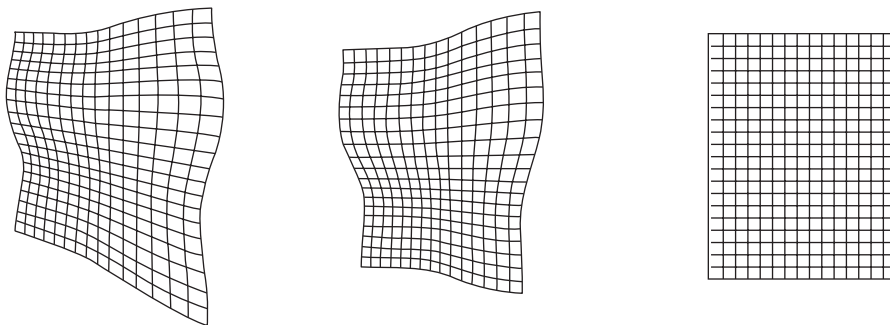
(a) Tilted Perspective Image

(b) Rectified Perspective Image
(Equivalent Vertical Photograph)

(c) Differentially Rectified Image
(Orthophotograph)

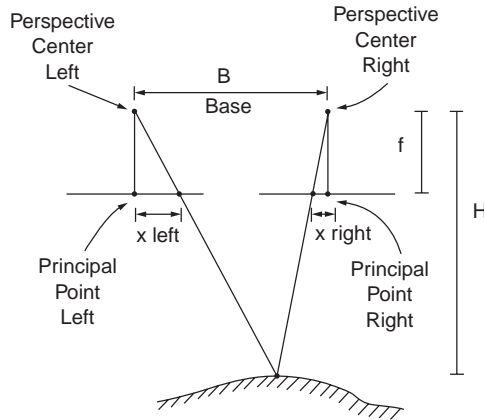**FIGURE 56.3** Image rectification sequence.

FIGURE 56.4 Parallax geometry.

Parallax and Stereo

Parallax is defined as the apparent shift in the position of an object, caused by a shift in the position of the viewer. Alternately closing one eye and then the other will demonstrate the concept of parallax, as near objects appear to shift, whereas far objects will appear stationary. This effect is the primary mechanism by which we achieve binocular depth perception, or stereo vision. This effect is exploited in photogrammetry when two overlapping photographs are taken. Within the overlap area, objects are imaged from two different exposure positions and parallax is evidenced in the resulting images. Parallax measurements can be used for approximate height computations in the following way using the geometry in Fig. 56.4. Two vertical overlapping photographs are taken at the same altitude, and an image coordinate system is set up within each image with the *x* axis parallel to the flight line, and with origin at the principal point in each image. For a given point, the parallax is defined as

$$p = x_{\text{left}} - x_{\text{right}}$$ (56.5)

The dimension *H* can be computed from

$$H = \frac{fB}{p}$$ (56.6)

B represents the base, or distance between the exposure stations. As is evident from Eq. (56.6) the distance *H* is inversely related to parallax, so that large parallax yields small *H*. Equation (56.6) is most often used with a pair of nearby points to determine a difference in elevation rather than the absolute elevation itself. This is given by the following equation, in which *b* is the image dimension of the base, and Δp is the difference in parallax.

$$\Delta H = \frac{H\Delta p}{b}$$ (56.7)

Figure 56.4 also illustrates the concept of *B/H* or *base–height ratio*. For a given focal length, format size, and overlap there is a corresponding *B/H* value for some average elevation in the area of interest. Large *B/H* yields strong geometry for elevation determination and small *B/H* yields weak geometry. For typical aerial photography for engineering mapping one would have *f* = 152 mm, *H* = 732 m, *B* = 439 m, overlap = 60%, and *B/H* = 0.6.

The viewing geometry as opposed to the taking geometry for this imagery yields a *B/H* of 0.3. The difference between the taking geometry and the viewing geometry causes *vertical exaggeration*.

That is, features appear stretched in the *z*-dimension compared to the planimetric dimension. This does not affect measurements, but only the viewer's perception of depth.

56.2   Sensors and Platforms

Cameras

Aerial cameras for engineering mapping use standard aerial film with 23-cm width. The emulsion type is usually panchromatic (black and white), but color film is becoming more popular. For some interpretation tasks, particularly involving vegetation, color infrared film is also used. This should not be confused with thermal infrared imaging, which requires a special sensor and cannot be directly captured on film. Resolution of camera components can be, individually, in the range of 140 line pairs per millimeter. But taken as a system the resolution would be in the range of 30 line pairs per millimeter. To maintain high resolution at low altitude, modern cameras employ image motion compensation to translate the film in synchronism with the apparent ground motion while the shutter is open. The ratio of image velocity in the image plane, v, to camera velocity over the ground, V, for a nominally vertical photograph is given by

$$\frac{v}{V} = \text{Scale} = \frac{f}{H} \tag{56.8}$$

A vacuum system holds the film firmly in contact with the platen during exposure, releasing after exposure for the film advance. During a photo flight, the camera is typically rotated to be parallel with the flight direction, rather than the aircraft body, in case there is a crosswind. This prevents *crab* between the adjacent photographs.

Calibration of aerial cameras should be executed every two to three years. Such a calibration determines the resolution, shutter characteristics, fiducial mark coordinates, principal point location, and a combination of focal length and lens distortion parameters that permits the modeling of the system to have central perspective geometry. Specifications for engineering mapping should always include a requirement for up-to-date camera calibration. A typical radial lens distortion curve is given in Fig. 56.5. The horizontal axis is radial distance from the principal point in millimeters, and the vertical axis is distortion at the image plane in micrometers.

Terrestrial cameras for photogrammetry usually employ a smaller film format than the standard aerial case; 70-mm film is a popular size for terrestrial photogrammetry. Whereas aerial cameras are at fixed, infinity focus, terrestrial cameras are typically used at a variety of focus positions. Thus calibration is more difficult since the lens components, and hence their image-forming attributes, are changing from one focus position to another. Often a camera manufacturer will provide principal distance and lens distortion data for each of several focus positions. Alternatively, the user may wish to model such behavior with additional parameters in a block adjustment with self-calibration.
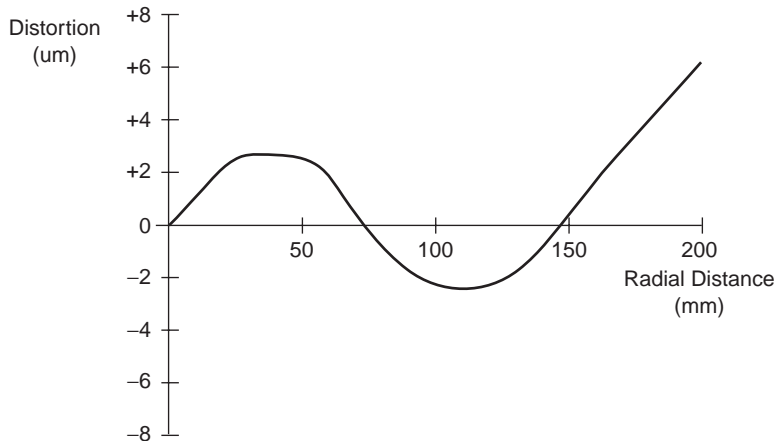

FIGURE 56.5 Lens distortion curve.

Before leaving the subject of photogrammetric cameras, mention should be made of the possibility of using an electronic image sensor in place of hardcopy film. In the case of large scale, aerial, engineering mapping photography, sensors of a size and resolution to duplicate the functionality of film are not available, and will not be for many years into the future. In the case of small format cameras, it is a different matter. Area sensors with 2048 × 2048 picture elements are available, with an element size of 10 micrometers. The total area of such an array would be therefore about 20 by 20 mm. This begins to approach the format size and resolution that could be usable for terrestrial photogrammetry. The direct capture of a digital image has advantages when feature extraction or enhancement by image processing is to be employed.

Scanners

Satellite imagers necessarily employ electronic sensors rather than film-based image capture because of the difficulty of retrieving film packets versus the relative ease of transmitting digital data by telemetry. A variety of sensor technologies is employed in satellite imagers, and the first of these, mechanical scanners, is discussed here. The MSS, multispectral scanner, on Landsats 1 to 5 is shown schematically in Fig. 56.6. For this instrument the mirror scans across the track of the satellite. The focal plane elements consist of light-gathering fiber-optic bundles

FIGURE 56.6 Schematic of mechanical scanner.

which carry the radiation to PMTs, photomultiplier tubes, or photodiode detectors. The *IFOV, instantaneous field of view,* determined by detector size, telescope magnification, and altitude, is 83 m. The scan rate and the sampling rate determine the *pixel size*. For the MSS it is 83 m in the along-track dimension and 68 m in the across-track dimension. The MSS has detectors in 4 bands covering the visible and near infrared portions of the spectrum. The TM, thematic mapper, on Landsats 4 and 5 has a similar design but enhanced performance compared to the MSS. Data recording takes place on both swings of the mirror, and there is a wider spectral range among the seven bands. The pixel size is 30 m at ground scale.
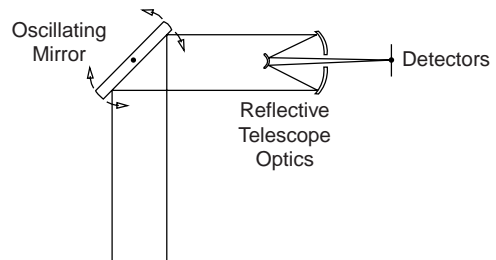
Pushbroom Linear Sensors

Linear CCDs, charge-coupled devices, are gaining wide acceptance for satellite imagers. A typical linear array is 2000 to 8000 elements in length. Longer effective pixel widths may be obtained by combining multiple arrays end to end (or equivalently using optical beam splitters). The individual elements in the sensor collect incident illumination during the integration period. The resulting charges are then transferred toward one end of the chip, and emerge as an analogue signal. This signal must be digitized for transmission or storage. Pixel sizes are usually in the 5- to 15-micrometer range. Since telescope optics have far superior resolution at the image plane, a purposeful defocusing of the optical system is performed to prevent aliasing. (See Fig. 56.7.)

56.3   Mathematics of Photogrammetry

Condition Equations

Using photogrammetry to solve spatial position problems inevitably leads to the formation of equations which link the observables to the quantities of interest. It is extremely rare that one would directly observe these quantities of interest. The form of the condition equations will reflect the nature of the observations, such as 2-D image coordinates or 3-D model coordinates, and the particular problem that one wishes to solve, such as space resection or relative orientation.
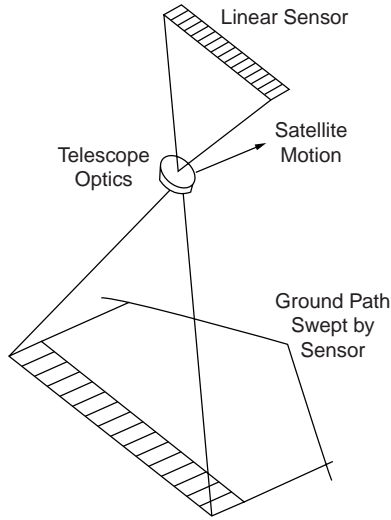
FIGURE 56.7 Pushbroom linear sensor.

Preliminaries

The 3×3 rotation that is often employed in developing photogrammetric condition equations is a function of three independent quantities. These quantities are usually the sequential rotations ω, ϕ, and κ, about the x, y, and z axes, respectively. The usual order of application is

$$M = M_\kappa M_\phi M_\omega \tag{56.9}$$

The elements are given by

$$M = \begin{bmatrix} \cos\phi\cos\kappa & \cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa & \sin\omega\sin\kappa - \cos\omega\sin\phi\cos\kappa \\ -\cos\phi\sin\kappa & \cos\omega\cos\kappa - \sin\omega\sin\phi\sin\kappa & \sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa \\ \sin\phi & -\sin\omega\cos\phi & \cos\omega\cos\phi \end{bmatrix} \tag{56.10}$$

An occasionally useful approximation to this matrix, in the case of small angles (i.e., near vertical imagery) is given by

$$M \approx \begin{bmatrix} 1 & \kappa & -\phi \\ -\kappa & 1 & \omega \\ \phi & -\omega & 1 \end{bmatrix} \tag{56.11}$$

in which the assumption is made that all cosines are 1, and the product of sines is zero. Other rotations and other parameters may also be used to define this matrix.

Collinearity Equations

The fundamental imaging characteristic of an ideal camera is that each object point, its corresponding image point, and the lens perspective center all lie along a line in space. This can be expressed in the following way, referring to Fig. 56.8:
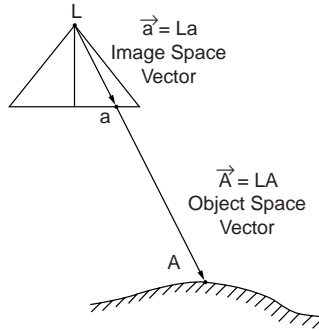
$$\vec{a} = k\ \vec{A} \tag{56.12}$$

FIGURE 56.8 Collinearity geometry.

This equation is valid only if the two vectors are expressed in the same coordinate system. The image space vector is inevitably expressed in an image coordinate system:

$$\vec{a} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} \qquad (56.13)$$

while the object space vector is expressed in an object space coordinate system (shifted to the camera perspective center):

$$\vec{A} = \begin{bmatrix} X - X_L \\ Y - Y_L \\ Z - Z_L \end{bmatrix} \qquad (56.14)$$

One thus needs to scale, rotate, and translate one of these two coordinate systems until they are coincident. This transformation is usually applied to the object space vector and is expressed as follows:

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = kM \begin{bmatrix} X - X_L \\ Y - Y_L \\ Z - Z_L \end{bmatrix} \qquad (56.15)$$

Eliminating the scale parameter k yields the classical form of the collinearity equations:

$$
\begin{aligned}
x - x_0 &= -f \left[\frac{m_{11}(X - X_L) + m_{12}(Y - Y_L) + m_{13}(Z - Z_L)}{m_{31}(X - X_L) + m_{32}(Y - Y_L) + m_{33}(Z - Z_L)} \right] \\
y - y_0 &= -f \left[\frac{m_{21}(X - X_L) + m_{22}(Y - Y_L) + m_{33}(Z - Z_L)}{m_{31}(X - X_L) + m_{32}(Y - Y_L) + m_{33}(Z - Z_L)} \right]
\end{aligned}
\qquad (56.16)
$$

Examples of particular problems for which the collinearity equations are useful include *space resection* (camera exterior orientation unknown, object points known, observed image coordinates given, usually implying a single image), *space intersection* (camera exterior orientations known, object point unknown, observed image coordinates given, usually implying a single object point), and *bundle block adjustment* (simultaneous resection and intersection, multiple images, and multiple points). This equation is non-linear and a linear approximation is usually made if we attempt to solve for any of the variables as unknowns. This dictates an iterative solution.
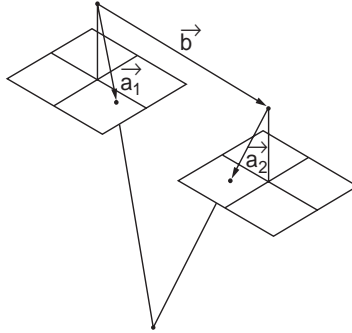
FIGURE 56.9 Coplanarity geometry.

Coplanarity Equation

The phrase *conjugate image points* refers to multiple image instances of the same object point. If we consider a pair of properly oriented images and the pair of rays defined by two conjugate image points, then this pair of rays together with the base vector between the perspective centers should define a plane in space. The *coplanarity* condition enforces this geometrical configuration. This is done by forcing these three vectors to be coplanar, which is in turn guaranteed by setting the triple scalar product to zero. An alternative explanation is that the parallelepiped defined by the three vectors as edges has zero volume. Figure 56.9 illustrates this geometry. The left vector is given by

$$a_1 = \begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = M_1^t \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix}_1 \tag{56.17}$$

The right vector is given by

$$a_2 = \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = M_2^t \begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix}_2 \tag{56.18}$$

and the base vector is given by

$$b = \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = \begin{bmatrix} X_{L_2} - X_{L_1} \\ Y_{L_2} - Y_{L_1} \\ Z_{L_2} - Z_{L_1} \end{bmatrix} \tag{56.19}$$

The coplanarity condition equation is the above-stated triple scalar product

$$F = \begin{vmatrix} b_x & b_y & b_z \\ u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{vmatrix} = 0 \tag{56.20}$$

The most prominent application for which the coplanarity equation is used is relative orientation. The equation is nonlinear and a linear approximation is usually made in order to solve for any of the variables as unknowns. This dictates an iterative solution.
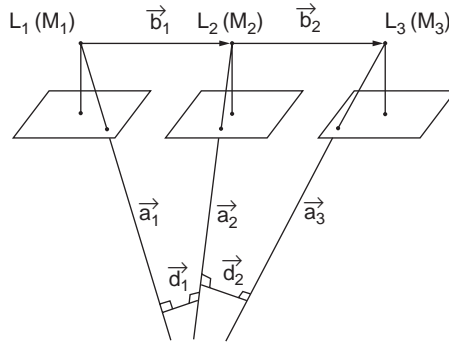
FIGURE 56.10 Scale restraint geometry.

Scale Restraint Equation

If photograph one is relatively oriented to photograph two, and photograph two is relatively oriented to photograph three, there is no guarantee that photograph one and photograph three are also relatively oriented. There are several methods to enforce this condition, and among them the most robust would be the *scale restraint* condition. This states that the intersection of conjugate rays from the three photographs should in fact occur at a single point. From Fig. 56.10 we see the three rays, $\mathbf{a}_i$, and two "mismatch" vectors, $\mathbf{d}_i$:

$$\mathbf{a}_1 = \begin{bmatrix} a_{1_x} \\ a_{1_y} \\ a_{1_z} \end{bmatrix} = \mathbf{M}_1^t \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \\ -f \end{bmatrix} \tag{56.21}$$

$$\mathbf{a}_2 = \begin{bmatrix} a_{2_x} \\ a_{2_y} \\ a_{2_z} \end{bmatrix} = \mathbf{M}_2^t \begin{bmatrix} x_2 - x_0 \\ y_2 - y_0 \\ -f \end{bmatrix} \tag{56.22}$$

$$\mathbf{a}_3 = \begin{bmatrix} a_{3_x} \\ a_{3_y} \\ a_{3_z} \end{bmatrix} = \mathbf{M}_3^t \begin{bmatrix} x_3 - x_0 \\ y_3 - y_0 \\ -f \end{bmatrix} \tag{56.23}$$

$$\mathbf{d}_1 = \mathbf{a}_1 \times \mathbf{a}_2 \tag{56.24}$$

$$\mathbf{d}_2 = \mathbf{a}_2 \times \mathbf{a}_3 \tag{56.25}$$

The scale restraint equation itself forces the independent scale factors for the common ray to be equal:

$$F = \frac{\begin{vmatrix} a_{1_x} & d_{1_x} & b_{1_x} \\ a_{1_y} & d_{1_y} & b_{1_y} \\ a_{1_z} & d_{1_z} & b_{1_z} \end{vmatrix}}{\begin{vmatrix} a_{1_x} & d_{1_x} & a_{2_x} \\ a_{1_y} & d_{1_y} & a_{2_y} \\ a_{1_z} & d_{1_z} & a_{2_z} \end{vmatrix}} + \frac{\begin{vmatrix} b_{2_x} & d_{2_x} & a_{3_x} \\ b_{2_y} & d_{2_y} & a_{3_y} \\ b_{2_z} & d_{2_z} & a_{3_z} \end{vmatrix}}{\begin{vmatrix} a_{2_x} & d_{2_x} & a_{3_x} \\ a_{2_y} & d_{2_y} & a_{3_y} \\ a_{2_z} & d_{2_z} & a_{3_z} \end{vmatrix}} = 0 \tag{56.26}$$

This equation is used primarily in the analytical formation of strips by successive relative orientation of image pairs in the strip. Common points in any photo triplet (that is, two adjacent models) would be subjected to the scale restraint condition. This equation is nonlinear and would require linear approximation for practical use in the given application.

Linear Feature Equations

It can happen, particularly in close-range photogrammetry, that the object space parameters of a straight line feature are to be determined. If at the same time stereo observation is either unavailable or difficult because of convergence or scale, then it becomes helpful if one can observe the feature monoscopically on each image without the need for conjugate image points. This can be elegantly accomplished with a condition equation which forces the ray associated with an observed image coordinate to pass through a straight line in object space. For each point in each photograph, a condition equation of the following kind may be written:

$$\begin{vmatrix} \rho_x & \rho_y & \rho_z \\ \beta_x & \beta_y & \beta_z \\ LC_x & LC_y & LC_z \end{vmatrix} = 0 \tag{56.27}$$

In this equation the vector ρ is the object space vector from the observed image point, the vector β is the vector along the straight line in object space, and LC is the vector from the perspective center to the point on the line closest to the origin. The six linear feature parameters must be augmented by two constraints which fix the magnitude of β to 1, and guarantee that β and C are orthogonal. A variation on this technique is the case where the object space feature is a circle in space. In this case each point on each image contributes an equation of the form

$$\left\| (L - C) - \frac{(L - C)\eta}{\rho\eta}\rho \right\| = r \tag{56.28}$$

where ρ and L have the same meaning as before. η represents the normal vector to the circle plane, r represents the circle radius, and C represents the circle center. In this case the normal vector must be constrained to unit magnitude. As in every case described here these equations are nonlinear in the variables of interest, and when we solve for them, the equations must be approximated using the Taylor series.

Block Adjustment

The internal geometry of a block of overlapping photographs or images may be sufficient to determine relative point positions, but for topographic mapping and feature extraction, one needs to tie this block to a terrestrial coordinate system. It would be possible to provide field survey determined coordinates for every point, but this would be prohibitively expensive. Thus arises the need to simultaneously tie all the photographs to each other, as well as to a sparse network of terrestrial control points. This process is referred to as *block adjustment*. The minimum amount of control necessary would be seven coordinate components, i.e., two horizontal (X, Y) points and three vertical (Z) points, or two complete control points (X, Y, Z) and one point with only vertical (Z). In practice, of course, one usually provides control in excess of the minimum, the redundancy providing increased confidence in the results.

Block Adjustment by Bundles

Block adjustment by bundles is the most mathematically rigorous way to perform this task. Observations consist of 2-D photograph image coordinates, usually read from a comparator or analytical plotter, transformed to the principal point origin, and refined for all known systematic errors. These systematic errors, described below, consist of at least lens distortion and atmospheric refraction, although at low altitude refraction may be considered negligible. Each image point, i, on each image, j, contributes two collinearity condition equations of the form
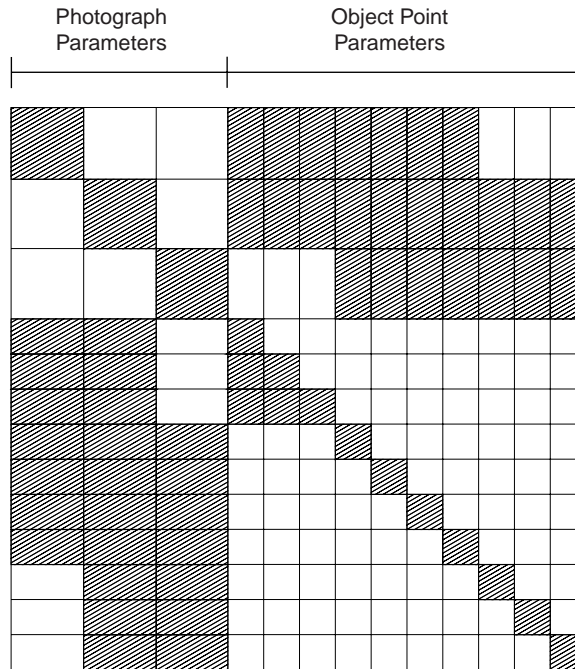
FIGURE 56.11 Organization of normal equations for bundle adjustment.

$$x_i - x_0 = -f \left[\frac{m_{11}(X_i - X_{L_j}) + m_{12}(Y_i - Y_{L_j}) + m_{13}(Z_i - Z_{L_j})}{m_{31}(X_i - X_{L_j}) + m_{32}(Y_i - Y_{L_j}) + m_{33}(Z_i - Z_{L_j})} \right]$$

$$\tag{56.29}$$

$$y_i - y_0 = -f \left[\frac{m_{21}(X_i - X_{L_j}) + m_{22}(Y_i - Y_{L_j}) + m_{33}(Z_i - Z_{L_j})}{m_{31}(X_i - X_{L_j}) + m_{32}(Y_i - Y_{L_j}) + m_{33}(Z_i - Z_{L_j})} \right]$$

in which (x_i, y_i) are the observed image coordinates, transformed into a coordinate system defined by the camera fiducial coordinates. The variables (x_0, y_0) represent the position of the principal point and f represents the focal length or principal distance. The last three variables would often be considered as fixed constants from a camera calibration report, or they may be carried as unknowns in the adjustment. The variables (X_i, Y_i, Z_i) represent the object coordinates of the point i. They may be known or partially known if point i is a control point, or they may be unknown if point i is a pass point. The variables $(X_{L_j}, Y_{L_j}, Z_{L_j})$ represent the coordinates of the exposure station or perspective center of image j. Each image, j, also has an associated orientation matrix, M_j, whose elements are shown in the equation. If there are n points observed on m images, the total number of condition equations will be $2n\,m$. The total number of unknowns will be $3n + 6m -$ (number of fixed coordinate components). With the advent of GPS in the photo aircraft, control may be introduced not only at the object points but also at the exposure stations. If the solution to the overdetermined problem is carried out by normal equations, the form of these equations is shown in Fig. 56.11.

Block Adjustment by Models

Block adjustment by models is necessary if the original observations of the photographs are made in stereo with an observed (x, y, z) model coordinate for each control point and pass point. Together with the model coordinates of the perspective centers, all model points are simultaneously transformed to the object (or control) coordinate system. There is usually redundant information to specify this transformation, so a
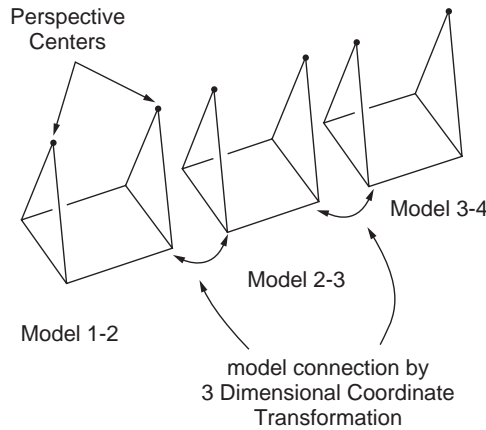
FIGURE 56.12 Model connection and strip formation.

least squares estimation is necessary. Practitioners have used comparator-derived image coordinates to compute model coordinates, and then subjected these derived model coordinates to an independent model block adjustment. This practice should be discouraged in favor of using the image coordinates directly in a bundle adjustment as described above. For the independent model block adjustment, each point, i, in model (stereo pair), j, will be related to the object space coordinates by a seven-parameter transformation unique to each model. The seven parameters include scale, s, rotations, Ω, Φ, and K, and translations T_x, T_y, T_z. For each point, i, in each model, j, the following three equations can be written:

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = -\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_i + s_j M_j \begin{bmatrix} x \\ y \\ z \end{bmatrix}_i + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}_j = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{56.30}$$

The uppercase coordinate vector represents the coordinate system of the control points, the lowercase vector represents the model coordinates, and the matrix M contains the rotation parameters. Only for control points will the (X, Y, Z) values be known; for all other points these will be unknown parameters solved for in the block adjustment. For n points and m models the total number of condition equations would be $3n\,m$. The total number of unknown parameters would be $3n + 7m -$ (number of fixed coordinate components).

Strip Formation and Block Adjustment by Polynomials

Strip formation and block adjustment by polynomials assumes that the input data are xyz model coordinates. These would usually come directly from a relatively oriented model in an analogue stereoplotter. They could also come from analytical computation from image coordinates. If this is the case it would be preferable to use the image coordinates directly in a bundle block adjustment. A similar comment was made with regard to the independent model block adjustment. The strip formation consists of linking successive models (including perspective center coordinates) by seven-parameter transformations, and then transforming each new model into the strip system based on the first model. This process is illustrated in Fig. 56.12. If a single strip is sufficiently short, say five models or less, the strip can be fitted to the control points by a global seven-parameter transformation. This is given in the following equation:

$$\begin{bmatrix} E \\ N \\ h \end{bmatrix} = sM \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \tag{56.31}$$

If the strip is longer than this, then because of adverse error propagation, artificial bends and bows will be present in the strip coordinates and polynomials present a way to model such effects. This technique is primarily of historical interest since its computational efficiency is no longer a compelling attribute. Both model coordinates and ground coordinates are transformed into an "axis of flight" coordinate system centered within the strip. Then conformal polynomials are used to transform the strip planimetric coordinates into the control system:

$$x' = x + a_1 + a_3x - a_4y + a_5(x^2 - y^2) - 2a_6xy + a_7(x^3 - 3xy^2) - a_8(3x^2y - y^2) + \cdots \quad (56.32)$$

$$y' = y + a_2 + a_4x + a_3y + a_6(x^2 - y^2) + 2a_5xy + a_7(3x^2y - y^3) + a_8(x^3 - 3x^2y) + \cdots \quad (56.33)$$

The vertical control points are used in the following polynomial:

$$z' = z + b_0 - 2b_2x + 2b_1y + c_1x^2 + c_2x^3 + c_3x^4 \\ + d_1xy + d_2x^y + d_3x^3y + d_4x^4y + e_1y^2 + e_2xy^2 \quad (56.34)$$

The number of terms is selected based on the quantity of control points and the length of the strip. Following the polynomial estimation, the points are transformed back into the original control coordinate system.

Image Coordinate Refinement

Raw stage coordinates from a comparator or analytical plotter must undergo a number of transformations and refinements before being used in further photogrammetric processing such as relative orientation or bundle block adjustment. Firstly the stage coordinates are transformed into the coordinate system defined by the camera *fiducial marks* or registration marks. This is usually done with a four- or six-parameter transformation. They are then shifted to the *principal point of autocollimation* by the principal point offsets (x_0, y_0). Following this they are corrected for radial lens distortion based on their position with respect to the *principal point of best symmetry*. The radial lens distortion is provided as part of the calibration of the camera either in the form of a table, a graph, or a polynomial function. A sample radial lens distortion graph is shown in Fig. 56.5. The usual form for polynomial lens distortion functions is given by

$$\Delta r = k_0r + k_1r^3 + k_2r^5 + k_3r^7 \quad (56.35)$$

in which the radial distance r is the distance from the symmetry point mentioned above. If a distortion table or function is given, the correction should be applied with the opposite sign. Conventionally, "+" indicates radial distortion outward from the principal point. Thus the correction equations would be

$$x_c = x\left(1 - \frac{\Delta r}{r}\right)$$
$$y_c = y\left(1 - \frac{\Delta r}{r}\right) \quad (56.36)$$

Following lens distortion correction the image coordinates should be corrected for atmospheric refraction (if it is significant, i.e., on the order of a micrometer or larger). The expression for radial image displacement due to atmospheric refraction is given by

$$d_r = K\left(r + \frac{r^3}{f^2}\right) \quad (56.37)$$

where r is the radial distance from the principal point of autocollimation and f is the focal length. The value for K is a function of the camera altitude and terrain elevation, and is given according to the ARDC Model Atmosphere (Air Research and Development Command of the U.S. Air Force):

$$K = \left[\frac{2410H}{H^2 - 6H + 250} - \frac{2410h}{h^2 - 6h + 250} \left(\frac{h}{H} \right) \right] \times 10^{-6} \qquad (56.38)$$

where H is the flying height in kilometers above sea level and h is the terrain height, also in kilometers above sea level. The displacement due to atmospheric refraction is always radially outward, therefore the correction is always radially inward. The same correction formulas may be used as in the case of lens distortion, replacing Δr by dr. Some practitioners have advocated handling earth curvature effects by modifying the image coordinates. This is to be discouraged. A better solution is to ensure that the object space coordinate system is truly Cartesian, and then the "problem" disappears. See the following section for a discussion of this.

Object Space Coordinate Systems

Geodetic Coordinates ϕ, λ, h

Geodetic coordinates, that is, latitude, longitude, and height, are the most fundamental way to represent the position of a point in space with respect to a terrestrial ellipsoid. However, photogrammetric condition equations are usually expressed in terms of rectangular, Cartesian coordinates. Thus, for the purpose of providing a reference frame for photogrammetric computations, one would usually transform ϕ, λ, and h into a rectangular system.

Space Rectangular Coordinates

Geocentric space rectangular coordinates may be derived from geodetic coordinates in such a way that the Z axis is parallel with the axis of rotation of the ellipsoid, and the X axis passes through the meridian of Greenwich in the plane of the equator. The Y axis is constructed so that the system is right-handed. In the following, h is assumed to be the ellipsoid height; if geoid height is given, it must be modified by the local geoid separation. The equations transforming geodetic coordinates into geocentric space rectangular coordinates are given by

$$\begin{aligned} X &= (N + h) \cos\phi \cos\lambda \\ Y &= (N + h) \cos\phi \sin\lambda \\ Z &= [N(1 - e^2) + h] \sin\phi \end{aligned} \qquad (56.39)$$

where N, the radius of curvature in the prime vertical is given by

$$N = \frac{a}{\sqrt{1 - e^2 \sin^2\phi}} \qquad (56.40)$$

where a = the semimajor axis of the ellipsoid
              b = the semiminor axis
              e = the eccentricity given by

$$e = \sqrt{\frac{(a^2 - b^2)}{a^2}} \qquad (56.41)$$

The inverse transformation cannot be given in a closed form. One can solve for ϕ, λ, and h by choosing an initial approximation and proceeding iteratively by the conventional Newton method,

$$X_{i+1} = X_i - J^{-1}F(X_i) \qquad (56.42)$$

or

$$
\begin{bmatrix} \phi_{i+1} \\ \lambda_{i+1} \\ h_{i+1} \end{bmatrix} = \begin{bmatrix} \phi_i \\ \lambda_i \\ h_i \end{bmatrix} - \begin{bmatrix} \dfrac{\partial F_1}{\partial \phi} & \dfrac{\partial F_1}{\partial \lambda} & \dfrac{\partial F_1}{\partial h} \\ \dfrac{\partial F_2}{\partial \phi} & \dfrac{\partial F_2}{\partial \lambda} & \dfrac{\partial F_2}{\partial h} \\ \dfrac{\partial F_3}{\partial \phi} & \dfrac{\partial F_3}{\partial \lambda} & \dfrac{\partial F_3}{\partial h} \end{bmatrix}^{-1} \begin{bmatrix} F_1(\phi_i, \lambda_i, h_i) \\ F_2(\phi_i, \lambda_i, h_i) \\ F_3(\phi_i, \lambda_i, h_i) \end{bmatrix}
\tag{56.43}
$$

and the three functions are the ones given in Eq. (56.39). One possible difficulty with the use of geocentric space rectangular coordinates is the large magnitude of the coordinate values. If one wished to maintain point precision to the nearest millimeter, 10 significant digits would have to be carried in the coordinates and single precision floating point computations would be insufficient. An alternative is the local space rectangular system, which is just the geocentric space rectangular coordinates, rotated so that Z' passes through a local point, Y' is in the meridian plane, and X' is constructed for a right-handed system. The LSR, or local space rectangular coordinates, are given by

$$
\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix}_{LSR} = M \left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{GSR} - \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \right)
\tag{56.44}
$$

where GSR refers to the geocentric space rectangular coordinate vector, the T vector is the translation to the local origin, and the rotation matrix M is given by

$$
M = \begin{bmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\sin\Phi\cos\lambda & -\sin\Phi\sin\lambda & \cos\Phi \\ \cos\Phi\cos\lambda & \cos\Phi\sin\lambda & \sin\Phi \end{bmatrix}
\tag{56.45}
$$

Map Projections Coordinates

The common map projections used to express terrestrial control points are the lambert conformal conic and the transverse mercator. In the U.S., each state has a state plane coordinate system utilizing possibly multiple zones of these projections. Globally, there is the UTM, or universal transverse mercator system, in which the globe is divided into 60 zones of width 6 degrees. Zone 1 is from 180 degrees west to 174 degrees west, and the zone numbering proceeds eastward until the globe is covered at zone 60. The zones are limited to ± 80 degrees latitude, and the scale factor at the central meridian is 0.9996. All of the above map projection coordinates have a common deficiency when used in photogrammetry. The XY coordinate is with respect to the developed projection surface, but the height coordinate is usually with respect to a sea level datum. Thus the system is not Cartesian. Over a very small region one could neglect the curved Z-reference, but over any substantial project area the nonorthogonality of the coordinate system will present itself in photogrammetric computations as a so-called earth curvature effect. The best approach to handling this situation is to either transform all control points into an LSR system described above or construct a local tangent plane system from the map projection coordinates, modifying the height component as follows:

$$
h_{tp} = h_{sl} - \frac{D^2}{2} R
\tag{56.46}
$$

where the subscript tp refers to the tangent plane system, sl refers to the sea level system, D is the distance from a project centered tangent point, and R is the nominal earth radius. Following all photogrammetric computations, that is, block adjustment, the heights can be corrected back into the sea level system for use by compilers and engineers.

56.4   Instruments and Equipment

Stereoscopes

Stereo viewing is possible with the unaided eyes if conjugate imagery is placed at a spacing approximately equal to the eye base, at a comfortable distance in front of the eyes. Prolonged viewing at such a distance may produce eye fatigue and therein lies the value of a stereoscope. A simple *lens stereoscope* allows the eyes to focus comfortably at infinity, thus permitting longer working sessions. For frame photographs, the overlapping pair should be laid out with the flight lines coincident and the spacing adjusted for comfortable viewing. Only a small portion of a standard 23-cm photograph overlap area can be viewed in this way, and some bending of the paper prints may be necessary to access the full model area. A *mirror stereoscope,* being larger, permits viewing of almost an entire overlap area, at a necessarily smaller scale. Approximate elevations can be read via a *parallax bar* and the associated 3-D measuring mark. Some modern softcopy stereo viewing systems employ nothing more than a simple mirror stereoscope to view conjugate imagery presented in split-screen mode on a video monitor.

Monocomparator, Stereocomparator, Point Marker

Both of the comparator instruments have been largely superseded by the analytical plotter, which is really nothing more than a computer-controlled stereocomparator. In any case, a monocomparator is a single two-axis stage with a measuring microscope and a coordinate readout, preferably with an accuracy of 1 or 2 micrometers. A stereocomparator is a pair of two-axis stages which permit stereo viewing by a pair of measuring microscopes, and simultaneous coordinate readout of two pairs of (XY) coordinates. Accuracy levels should be comparable to that mentioned for the monocomparator. Both of these comparator instruments are used chiefly for *aerial triangulation, bridging,* or *control extension.* In this process, all *control points* and *pass points* are read for all photographs in a *strip* or *block.* The photos are then linked by geometric condition equations and tied to the ground coordinate system, thus producing ground coordinates for all observed pass points. These pass points may then be used for individual model setups in a stereo restitution instrument. If pass points are desired in an area of the photograph without identifiable detail points, artificial emulsion marks or "pug points" are introduced by a *point marker* or "pug." These marks are typically 40- to 80-micrometer-diameter drill holes in the photograph emulsion, sized to be compatible with the stereo measuring device.

Stereo Restitution: Analogue, Analytical, Softcopy

Early instruments for map compilation consisted of optical projectors and a small viewing screen with a means to direct the image from one projector to the left eye and from the other projector to the right eye. This binocular separation was effected by *anaglyph* (red and blue filters), by mechanical shutter, and by polarization. Analogue instruments in use today employ exclusively mechanical projection in which a collection of gimbals, space rods, and cardan joints emulate the optical light paths. All analogue instruments must provide a way to re-create the inner camera geometry by positioning the principal point (via the fiducial marks) and setting the principal distance or focal length. These steps constitute the interior orientation. A procedure is also necessary to reestablish the *relative orientation* of the photographs at instant of exposure. This is accomplished by clearing *y-parallax,* or *y* displacement in model space between the projected images, in at least five points spaced throughout the model. For the point layout in Fig. 56.13 the sequence of steps for two-projector relative orientation is as follows:

1. Clear at point 1 with kappa-right.
2. Clear at point 2 with kappa-left.
3. Clear at point 3 with phi-right.
4. Clear at point 4 with phi-left.
5. Clear at point 5 with omega-left or omega-right.
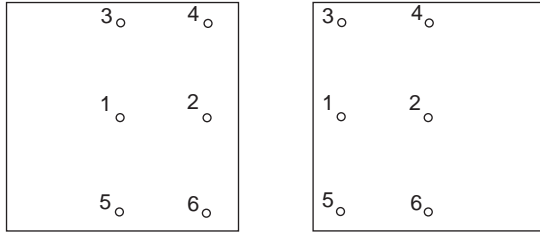6. Check for no parallax at point 6.

FIGURE 56.13 Point layout for relative orientation.

If there is parallax at point 6, the procedure is repeated until no parallax is seen at point 6. Convergence can be speeded up by overcorrecting by about half at step 5. When this is complete, the entire model should be free of y-parallax. If there is visible parallax at other points, it could be due to uncompensated lens distortion, excessive film deformation, or other factors. Following relative orientation comes the *absolute orientation,* in which the relation is established between the model coordinates and the ground coordinates, defined by control points in the model. In the past this was done by physically orienting a map manuscript to a mechanical tracing device. This physical procedure would involve scaling, by adjusting the base components, and leveling, by adjusting either common rotation elements or combinations of projector rotations and corresponding base components. Now it is done analytically by computing the parameters of the three-dimensional similarity transformation between the model and ground coordinates. The computed rotations would then be introduced into the instrument as before. This computationally assisted absolute orientation requires that the instrument be fitted with position encoders for coordinate readout of *xyz* model coordinates. Accuracies on the order of 5 micrometers are typically seen for this task. Schematic depictions of an optical and a mechanical stereo restitution instrument are shown in Fig. 56.14(a) and (b). In addition to map compilation of planimetry and elevation data, an analogue stereo instrument can also be used to collect model coordinates for independent model aerial triangulation. This requires an additional step of determining the model coordinates of the perspective center, which is necessary to link adjacent models in a strip.

All of the functions of an analogue instrument can be duplicated and usually exceeded in an *analytical plotter*. Such a device, shown schematically in Fig. 56.14(c), consists of two computer-controlled stages,
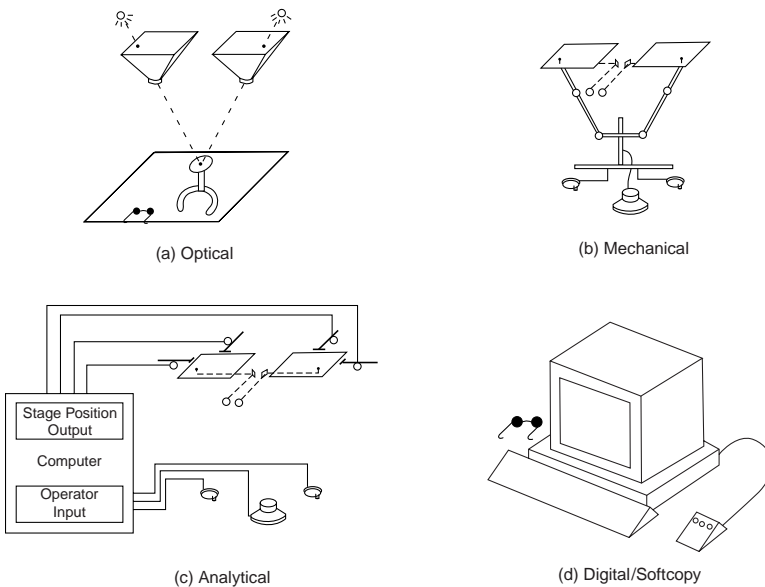

(a) Optical

(b) Mechanical

(c) Analytical

(d) Digital/Softcopy

FIGURE 56.14 Schematic diagrams of stereo restitution instruments.

a viewing stereomicroscope, operator controls for three-axis motion, and a suite of computer software to automate and assist in all of the desired operations. Interior orientation consists of measuring the fiducial marks and introducing the calibrated camera parameters. Relative orientation consists of measuring conjugate points and computing the five orientation parameters. Absolute orientation consists of measuring the control points and computing the seven-parameter transformation as above. Of course, the two steps of relative and absolute orientation can be combined in a two-photo bundle solution using the collinearity equations as described previously. In addition to conventional map compilation, analytical plotters are well suited to aerial triangulation and block adjustment, digital elevation model collection, cross section and profile collection, and terrestrial or close-range applications. Stage accuracies are typically 1 or 2 micrometers. Today one would always have a CAD system connected to the instrument for direct digitizing of features into the topographic or GIS database.

The most recent variant on the stereo viewer/plotter is the *softcopy* stereo system. Here the stereo images are presented to the operator on a computer video monitor. This is shown schematically in Fig. 56.14(d). In the two previous cases the input materials were hardcopy film transparencies. In this case the input material is a pair of digital image files. These usually come from digitized photographs, but can also come from sensors which provide digital image data directly, such as SPOT. Softcopy stereo systems present interesting comparisons with hardcopy-based instruments. Spatial resolution may be inferior to that visible in the original hardcopy image, depending on the resolution and performance of the scanner, but possibilities for automation and operator assistance by digital image processing are abundant and are being realized today. In addition, the complicated task of overlaying vector graphics onto the stereo images in a hardcopy instrument becomes a simple task in a softcopy environment. This can be enormously beneficial for editing and completeness checking. The orientation aspect of a softcopy system is very similar to the analytical plotter in that all computations for orientation parameters are done via computer from image coordinate input. The dramatic impact of softcopy systems will not be apparent until specialized image processing tools for feature extraction and height determination are improved to the point that they can reliably replace manual operation for substantial portions of the map compilation task. A few definitions are now presented to encourage standardized terminology. *Digital mapping* refers to the collection of digital map data into a CAD or GIS system. This can be done from any type of stereo device: analogue, analytical, or softcopy. *Digital photogrammetry* refers to any photogrammetric operations performed on digital images. *Softcopy photogrammetry* is really synonymous with digital photogrammetry, with the added connotation of softcopy stereo viewing.

Scanners

With the coming importance of digital photogrammetry, scanners will play an important role in the conversion of hardcopy photograph transparencies into digital form. For aircraft platforms, and therefore for the majority of large scale mapping applications, film-based imaging is still preferred because of the high resolution and the straightforward geometry of non-time-dependent imagery. Thus arises the need for scanning equipment to make this hardcopy imagery available to digital photogrammetric workstations. To really capture all of the information in high-performance film cameras, a pixel size of about 5 micrometers would be needed. Because of large resulting file sizes, many users are settling for pixel sizes of 12 to 30 micrometers. For digital orthophotos, sometimes an even larger size is used. Table 56.1 shows the relation between pixel size and file size for a 230-mm square image assuming no compression and assuming that each pixel is quantized to one of 256 gray levels (8 bits).

There are three main scanner architectures: (1) drum with a point sensor, usually a PMT (photomultiplier tube); (2) flatbed with area sensor; and (3) flatbed with linear sensor. Radiometric response of a scanner is usually set so that the imagery uses as much of the 256 level gray scale as possible. The relation between gray level and image density should be known from system calibration. In some cases, gray values can be remapped so that they are linear with density or transmittance. Most photogrammetric scanners produce color by three passes over the imagery with appropriate color filters. These can be recorded in a band sequential or band interleaved manner as desired. There are a large number of image file formats in use.

TABLE 56.1   File Sizes for Given Pixel Sizes

Pixel Size	File Size, 230-mm Image
5 μm	2.1 Gb
10 μm	530 Mb
15 μm	235 Mb
20 μm	130 Mb
25 μm	85 Mb
50 μm	21 Mb
100 μm	5 Mb

Plotters

Until recently the majority of engineering mapping (as opposed to mass production mapping) has been produced on vector plotters. These produce vector line work on stable base material. They are usually based on a rotating drum for one axis and a moving pen carriage for the other axis. Flatbed designs also exist with a two-axis cross-slide for the pen carriage. These have the additional possibility to handle scribing directly in addition to ink. Electrostatic plotters are essentially raster plotters which may emulate a vector plotter by vector to raster conversion. Digital photogrammetry, with the integration of images and vectors, requires a raster-oriented device. Likewise GIS, which often calls for graphic presentations with area fills or raster layers, may require a raster device.

56.5   Photogrammetric Products

Topographic Maps

The classical product of photogrammetric compilation is a *topographic map*. A topographic map consists typically of planimetric features such as roads, buildings, and waterways, as well as terrain elevation information usually in the form of contours. In the past these were manually drafted in ink or scribed onto scribe coat material. Today they are recorded directly into a CAD system or GIS. U.S. National Map Accuracy Standards (NMAS) dictate an accuracy of planimetric features as well as contour lines, which is tied to hard copy scale. In the digital environment, such standards may need to be revised to reflect the increasingly prominent role of the (scaleless) digital map representation. The following is a summary of the Office of Management and Budget standards:

1. Horizontal accuracy. For maps with publication scale greater than 1:20,000, not more than 10% of the "well-defined" points tested shall be in error by more than 1/30th of an inch at publication scale. For maps with publication scale less than 1:20,000, the corresponding tolerance is 1/50th of an inch.
2. Vertical accuracy. Not more than 10% of the elevations tested shall be in error by more than one-half contour interval. Allowances are made for contour line position errors as above.
3. Any testing of map accuracy should be done by survey systems of a higher order of accuracy than that used for the map compilation.
4. Published maps meeting these accuracy requirements shall note this fact in their legends, as follows: "This map complies with national map accuracy standards."
5. Published maps whose errors exceed these limits shall omit from their legends any mention of compliance with accuracy standards.
6. When a published map is a considerable enlargement of a map designed for smaller-scale publication, this fact shall be stated in the legend. For example, "This map is an enlargement of a 1:20,000 scale map."

Other commonly accepted accuracy standards are as follows. Reference grid lines and control point positions should be within 1/100th of an inch of their true position. Ninety percent of spot elevations

should be accurate to within one-fourth contour interval, and the remaining 10% shall not be in error by more than one-half contour interval.

Image Products

Image products from photogrammetry include uncontrolled mosaics, controlled mosaics, rectified enlargements, and orthophotos. Mosaics are collections of adjoining photograph enlargements mated in a way to render the join lines as invisible as possible. In the case of controlled mosaics, the photographs are enlarged in a rectifier, using control points to remove the effects of camera tilt, and to bring all enlargements for the mosaic to a common scale. Uncontrolled mosaics are similar except that no tilt removal is done, enlargement scales are less accurately produced, and continuity is attempted by the careful matching of image features. In the past all mosaicking has been done with paper prints, glue, and considerable manual dexterity. If the photographs are scanned, or if the imagery is originally digital, then the mosaicking process can be entirely digital. Digital techniques allow great flexibility for such tasks as tone matching, scaling, rectifying, and vector/annotation addition. Orthophotos are photographs which have been differentially rectified to remove both tilt displacements as well as relief displacements. A well-produced orthophoto can meet horizontal map accuracy standards, and can function as a planimetric map. Individual orthophotos can be further merged into orthophoto mosaics. Digital techniques for orthophoto production are also becoming very popular because of the flexibility mentioned above. People who are not mapping specialists seem to have a particularly easy time interpreting an orthophoto, compared to an abstract map with point, line, and area symbology that may be unfamiliar. With digital orthophoto generation, it is particularly effective to overlay contour lines on the imagery.

Digital Elevation Models

The concept of digitally recording discrete height points to characterize the topographic surface has been in practice for a number of years. Names and acronyms used to describe this concept include *DTM, digital terrain model; DEM, digital elevation model; DHM, digital height model;* and *DTED, digital terrain elevation data.* The philosophy behind this concept is that one obtains sufficient digital data to describe the terrain, and then generates graphic products such as contours or profiles only as a means for visualizing the terrain. This is in contrast to the conventional practice of recording a contour map and having the contours be the archival record which represents the landforms. The advantage of the DEM approach is that the height database can be used for several different applications such as contour generation, profile/cross section generation, automated road design, and orthophoto interpolation control. Potential pitfalls in this approach mostly revolve around decisions to balance the conflicting demands of accurate terrain representation versus fast data collection and reasonable file sizes. There are basically two alternatives to consider when collecting such data: random data points selected to describe the terrain, or a regular grid of points, with interval selected to describe the terrain.

Random Data Points

Random data points in a DEM may be used directly in a *TIN, triangulated irregular network*, or they may be used to interpolate a regular grid via a variety of interpolation methods. The TIN may be created by a number of algorithms which are producing the equivalent *Dirichlet tessellation, Thiesson Polygons,* or the *Delauney triangulation.* One of the simpler methods is the basic Watson algorithm, described by the following steps:

1. Create three fictitious points such that the defined triangle includes all of the data points.
2. Pick a point.
3. Find all of the triangles whose "circumcircle" (the circle passing through triangle vertices) contains the point.
4. The union of all triangles in step 3 forms an "insertion polygon."
5. Destroy all internal edges in the insertion polygon, and connect the current point with all vertices of the polygon.

6. Go to step 2, until no more points are left.
7. When done, eliminate any triangle with a vertex consisting of one of the initial three fictitious points.

To enforce a breakline, one can overlay the breakline on the preliminary TIN, and introduce new triangles as required by the breakline. Interpolation within a TIN usually means locating the required triangular plane facet, and evaluating the plane for Z as a function of X and Y. Contour line generation is particularly easy within the triangular plane facets, with all lines being straight and parallel. Connecting contour lines between facets requires a searching and concatenation operation.

Interpolation in a random point DEM not organized as a TIN can be carried out by various *moving surface* methods, or by *linear prediction*. An example of a moving surface model would be a moving "tilted" plane:

$$z = a_0 + a_1 x + a_2 y \tag{56.47}$$

One equation is written for each point within a certain radius, possibly with weighting inversely related to distance from desired position, and the three parameters are solved for, thereby allowing an estimate of the interpolated height. Such moving surface models can be higher-order polynomials in two dimensions, as dictated by point density and terrain character. With linear prediction, an elevation is interpolated as follows:

$$z_0 = \sigma^t_{zz_0} \Sigma^{-1}_{zz} z \tag{56.48}$$

where z is an $n \times 1$ vector representing height points in the vicinity of the point to be interpolated, Σ_{zz} is an $n \times n$ matrix of covariances between the reference points, usually based on distance, and σ_{zz_0} is an $n \times 1$ vector representing the covariances between the point to be interpolated and the reference points, again based on distance. Breaklines can be enforced in these methods by not allowing reference points on one side of the break to influence interpolations on the other side of the break.

Gridded Data Points

Points can be easily collected directly in a regular grid by programming an analytical stereo instrument to move in this pattern. Grids may also be "created" by interpolating at grid "posts" using random data as outlined above. Interpolation within a regular grid could be done by the methods outlined for random points, but is more often done by bilinear interpolation. Bilinear interpolation can be done by making two linear interpolations along one axis, followed by a single linear interpolation along the other axis. Alternatively, one could solve uniquely for the following four parameters a, b, c, d at the grid cell corners, and evaluate the equation at the unknown point. x and y can be local grid cell coordinates.

$$z = a + bx + cy + dxy \tag{56.49}$$

Contouring is also relatively straightforward in a gridded DEM by linear interpolation along the edges of each grid cell. An ambiguity may arise when there is the same elevation on all four sides of the grid cell.

Geographic Information Systems and Photogrammetry

Historically, map data collected over the same region for different purposes was typically stored separately with no attempts at registration and integration. Likewise, textual attribute data describing the map features were likely kept in yet another storage location. Increasingly, the trend, particularly for municipalities, is to coregister all of this diverse map and attribute data within a *geographic information system,* or GIS. The photogrammetric process typically plays a very important role in creating the *land base,* or *land fabric,* based on well-defined geodetic control points, to which all other GIS data layers are registered. Photogrammetry also plays a role in making periodic updates to the GIS land base in order to keep it current with new land development, subdivision, or construction. The photogrammetric compiler may

also be involved in collecting facilities features and tagging them with attributes or linking them to existing attribute records in a facilities database.

56.6   Digital Photogrammetry

The term *digital photogrammetry* refers to photogrammetric techniques applied to digital imagery. The current trend in favor of digital photogrammetry has been driven by the enormous increase in computer power and availability, and the advances in image capture and display techniques.

Data Sources

Digital image data for photogrammetry can come from a number of sources. Satellite imagers transmit data directly in digital form. Primary among the satellite sources would be SPOT and Landsat, each with a series of spacecraft and sensors to provide continuous availability of imagery. Digital cameras for airborne or terrestrial use have up to now been a minor contributor as source imagery for photogrammetry. This is principally due to the limited size and density of sensor element packing for area sensors, and to the adverse restitution capabilities with linear sensors using motion-induced scanning. The predominant source of digital imagery for engineering photogrammetry is the scan conversion of conventional film transparencies on high-accuracy scanning systems. Such image data is typically collected as 8-bit (256 level) gray values for monochrome imagery, or as 3×8-bit red, green, and blue components for color imagery. Monochrome digital images are stored as a *raster,* or rectangular grid format. For fast access they may also be *tiled,* or subdivided into subgrids. Color images may be stored in *band sequential* or *band interleaved* order.

Digital Image Processing Fundamentals

Sampling

Digital image data is generated by sampling from a continuous source of incident radiation. In the case of direct imaging, the source is truly continuous; in the case of sampling from a silver-halide emulsion image, the source is usually of sufficiently higher resolution as to be effectively continuous. The *sampling theorem* states that for a band-limited spatial signal with period T corresponding to the highest frequency present, we must sample it with a period no greater than $T/2$ in order to be able to reconstruct it without error. If there are frequencies present which are higher than can be reconstructed, *aliasing* will occur and may be visible in the reconstructed image as *moiré patterns* or other effects. If this is the case, one should purposely blur or defocus the image until the maximum frequencies present are consistent with the sampling interval. This creates an antialiasing filter. The sampled image will be referred to as $F(x, y)$. The gray values recorded in the digital image may represent estimates of well-defined photometric quantities such as irradiance, density, or transmittance. They may also be a rather arbitrary quantization of the range of irradiance levels reaching the sensor.

Histogram Analysis and Modification

The *histogram* of a digital image is a series of counts of frequency of occurrence of individual gray levels or ranges of gray levels. It corresponds to the probability function or probability density function of a random variable. Any operation on an image, A,

$$B(x, y) = f[A(x, y)] \tag{56.50}$$

in which the output gray level at a point is a function strictly of the input gray level is referred to as a *point operation* and will result in a modification of the histogram. If an image operation is a function of the position as well as the input gray level,

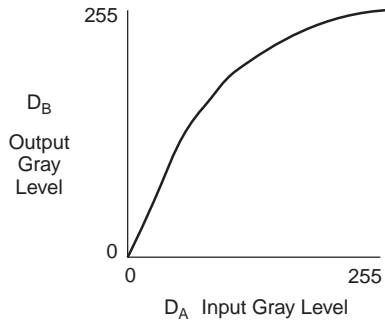$$B(x, y) = f[A(x, y), x, y] \tag{56.51}$$

FIGURE 56.15 Gray level remapping.

This may be referred to as a *spatially variant point operation*. It will result in a modification of the histogram that is position dependent. This would be used to correct for a sensor which has a position-dependent response. To illustrate a point operation, function $f(D)$ in Fig. 56.15 maps an input gray level D_A to an output gray level D_B. We assume for simplicity that f is monotone increasing or decreasing, and therefore has an inverse. The output histogram, $H_B(D)$, as a function of the input histogram, $H_A(D)$, and of the function, f, is

$$H_B(D) = \frac{H_A[f^{-1}(D)]}{|df/dD|} \tag{56.52}$$

This is analogous to the transformation of distributions based on a functional relation between random variables. Other common applications of histogram modification include contrast stretching and brightness shifting, that is, gain and offset.

Resampling, Geometric Image Manipulation

Whenever the geometric form of an image requires modification, *resampling* is necessary. This is really just two-dimensional interpolation. Examples of such geometric modifications are rotation, rectification, and epipolar line extraction. If a simple change of scale is desired, magnification can be accomplished by *pixel replication,* and minification can be accomplished by *pixel aggregation,* or more quickly (with aliasing) by *subsampling*. For the resampling task, the fastest method is *nearest neighbor,* which assigns the gray level of the nearest pixel center to the point being interpolated. The most common method would be *bilinear* interpolation. This math model was described under digital elevation models. Less common would be interpolation by higher-order polynomials, such as cubic polynomials. As an example, the process to generate a digital orthophoto can be easily summarized in this context. A grid is defined in object space, which represents the locations of the "output" pixels. The elevation for these points is observed directly or interpolated from a digital terrain model. Each of these (XYZ) object points is passed through the collinearity equations using the known orientation of an image, and image coordinates are obtained. These image coordinates are transformed into the row/column coordinate space of the digital image. If the output spacing is on the order of the digital image pixel spacing, bilinear interpolation can be used to interpolate the orthophoto pixel gray level. If the output spacing is much greater than the image spacing, the output gray value should be obtained by averaging over all image pixels within the larger pixel coverage, thus avoiding aliasing problems.

Filtering

For continuous functions in one dimension, *convolution* is expressed as

$$y(t) = f * g = \int_{-\infty}^{\infty} f(\tau)\, g(t - \tau)\, d\tau \tag{56.53}$$

For continuous functions in two dimensions, convolution is given by

$$h(x, y) = f * g = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v)g(x - u, y - v)\ du\ dv \tag{56.54}$$

For discrete two-dimensional functions, that is, images, convolution is given by

$$H(i, j) = F * G = \sum_m \sum_n F(m, n)G(i - m, j - n) \tag{56.55}$$

Image enhancements and feature exaggeration can be introduced by convolving the following 3×3 kernels with an image. For an edge parallel to azimuth 90 degrees,

$$G = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix} \tag{56.56}$$

For an edge parallel to azimuth 135 degrees,

$$G = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix} \tag{56.57}$$

For general edge sharpening via a Laplacian,

$$G = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{56.58}$$

For an edge parallel to azimuth 0 degrees using a Sobel operator,

$$G = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \tag{56.59}$$

For low pass filtering,

$$G = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \tag{56.60}$$

Matching Techniques

Digital image matching represents one of the means by which digital photogrammetry can potentially far exceed the productivity of conventional photogrammetry. Point mensuration and stereo compilation are tasks requiring skilled operators if done manually, and few means are available to speed up the manual process. Point mensuration and stereo "perception" by image matching will increase in speed with each advance in computer technology, including parallel processing. Point mensuration occurs at numerous stages in the photogrammetric restitution process, such as interior orientation (fiducial marks), relative

orientation (parallax points), absolute orientation (signalized control points), and aerial triangulation (pass points and signalized control points). In the case of signalized points, a strategy would be as follows:

1. Derive a rotation-independent detection template for the target type. Convolve this template with the image(s) under study, and, by thresholding, obtain the approximate locations of the points.
2. With a fine-pointing template estimate the target position in the image, while simultaneously modeling rotation, scale, affinity, and radiometry.

The search criterion can be the maximization of a correlation coefficient, or the minimization of a sum of squares of residuals. The correlation coefficient is given by

$$C_{uv} = \frac{\Sigma(u_i - \bar{u})(v_i - \bar{v})}{[\Sigma(u_i - \bar{u})^2 \Sigma(v_i - \bar{v})^2]^{1/2}} \qquad (56.61)$$

where u and v represent image and template, or vice versa, and

$$\bar{u} = \Sigma\frac{u_i}{N} \qquad \text{and} \qquad \bar{v} = \Sigma\frac{v_i}{N} \qquad (56.62)$$

The other significant application area of digital image matching is digital elevation model extraction. Three techniques will be described for this task: (1) vertical line locus, (2) least squares matching, and (3) epipolar matching.

Vertical Line Locus

Vertical line locus (VLL) is used to estimate the elevation at a single point, appearing in two photographs. An initial estimate is required of the object space elevation of the point to be estimated. A "search range" is then established in the Z-dimension, extending above and below the initial estimate. This search range is then subdivided into "test levels." At each test level, a matrix of points is defined surrounding the point to be estimated, all in a horizontal plane. All points in this matrix are then projected back into each of the two photographs, and via interpolation in the digital image, a corresponding matrix of gray levels is determined for each photograph, at each level. When the test level most nearly coincides with the actual terrain elevation at the point, the match between the pair of gray level matrices should be the maximum. Some measure of this match is computed, often the correlation coefficient, and the elevation corresponding to the peak in the match function is the estimated elevation. Variants on this procedure involving an iterative strategy of progressively finer elevation intervals, variable sized matrix, and "matrix shaping" based on estimated terrain slope can all be implemented, yielding more accurate results at the expense of more computing effort.

Least Squares Matching

In the usual least squares matching (LSM) approach, one assumes that the two images are the same except for an affine geometric relationship and a radiometric relationship with a gain and offset. One takes the gray values from one image of the pair to be observations and then computes geometric and radiometric transformation parameters in order to minimize the sum of squares of the discrepancies (residuals) with the second image. In the following equations, g will represent the first image and h will represent the second image. The affine geometric relationship is defined by

$$x_h = a_1 x_g + a_2 y_g + a_3 \qquad (56.63)$$

$$y_h = b_1 x_g + b_2 y_g + b_3$$

The linearized equation relating the gray levels is

$$g(x, y) = h(x, y) + \frac{\partial h}{\partial x}dx + \frac{\partial h}{\partial y}dy + h(x, y)\, dk_1 + dk_2 \qquad (56.64)$$

Taking differentials of Eq. (56.63),

$$dx_h = x_g\,da_1 + y_g\,da_2 + da_3 \qquad (56.65)$$

$$dy_h = x_g\,db_1 + y_g\,db_2 + db_3$$

making the following substitution for compact notation:

$$h_x = \frac{\partial h(x,y)}{\partial x} \qquad \text{and} \qquad h_y = \frac{\partial h(x,y)}{\partial y} \qquad (56.66)$$

adding a residual to g, and substituting the differentials into Eq. (56.64), we obtain the condition equation to be used for the least squares estimation,

$$g + v_g - h - h_x x_g\,da_1 - h_x y_g\,da_2 - h_x\,da_3 - h_y x_g\,db_1 \\ - h_y y_g\,db_2 - h_y\,db_3 - h\,dk_1 - dk_2 = 0 \qquad (56.67)$$

In matrix form,

$$v + B\Delta = f \qquad (56.68)$$

the condition equation becomes

$$v_g + \begin{bmatrix} -h_x x_g & -h_x y_g & -h_x & -h_y x_g & -h_y y_g & -h_y & -h & -1 \end{bmatrix} \begin{bmatrix} da_1 \\ da_2 \\ da_3 \\ db_1 \\ db_2 \\ db_3 \\ dk_1 \\ dk_2 \end{bmatrix} = h - g \qquad (56.69)$$

One such equation may be written for each pixel in the area surrounding the point to be matched. The usual solution of the least squares problem yields the parameter estimates, in particular the shift parameters to yield the "conjugate" image point. If the parameters are not small, it may be necessary to resample for a new "shaped" image, h, and solve repeated iterations until the parameter estimates are sufficiently small.

Epipolar Matching

Both of the previous techniques involved matching areas, albeit small ones. In epipolar matching, the two matched signals are one-dimensional. Also of interest is the degree to which these methods make use of a priori knowledge of the image orientation. In the VLL technique we make use of this information. In LSM we do not. In epipolar matching this information is used.

In theory, the use of this information should further restrict the solution space and thereby yield faster solutions. In the epipolar technique, one takes the two planes defined by the two images to be matched, and a third plane defined by the two perspective centers and intersecting the two photograph planes. This third plane is referred to as an epipolar plane (there is an infinite number of them). The intersection of this epipolar plane with each of the photograph planes defines two lines, one in each of the photographs. If the orientation is correct, it is guaranteed that each point on one of the lines has a conjugate point on the other line. Thus to search for this conjugate point one needs only to search in one dimension. The practical steps necessary to implement this technique would be as follows:

1. For each photograph, determine two points in object space which lie in the epipolar plane.
2. For each photograph, project the two corresponding points into the photograph plane and solve for the parameters defining the resulting line in image space.
3. Resample each digital image at an appropriate interval along these lines.
4. Determine a match point interval in image space. Select points in one image at this spacing along the epipolar line.
5. For each match point in one image, bracket it by some pixels on either side of it. Then search the epipolar line in the other image for the best match "element." The match criterion is usually the above-mentioned cross-correlation function, but it can be any objective function which measures the degree of "sameness" between the elements.
6. Create in this way an irregularly spaced line of points in object space.
7. Rotate the epipolar plane about the photograph baseline by a small amount, and repeat the process. This will create an irregular grid of *XYZ* points in object space.
8. Use these points directly to form a TIN, or interpolate a regular grid as described earlier.

Up to now, digital image matching has been most effective and reliable when used on small and medium scale imagery. Large scale images, showing the fine structure of the terrain along with individual trees, buildings, and other man-made objects, may contain steep slopes and vertical planes (i.e., not a smooth, continuous surface). This generally interferes with the simple matching algorithms described here. For large scale engineering mapping, more research needs to be done to develop more robust matching methods.

56.7 Photogrammetric Project Planning

Project planning usually starts with analysis of the specifications for the maps or image products to be produced. Requirements for these final products generally place fairly rigid constraints on the choices available to the planner. The constraints arise because the specified accuracies must be achieved and on the other hand the process must be economical.

Flight Planning

For conventional topographic mapping in the U.S., a compilation system (camera, field survey system, stereoplotter, operators, etc.) is often characterized by a *C-factor*, used as follows:

$$\text{Flying height} = C\text{-factor} \times \text{Contour interval} \tag{56.70}$$

This *C*-factor can be in the range of 150 to 250 depending on the quality of the equipment and the skill of the personnel doing the compilation. Thus if one knows from the specifications that a 1-m contour interval is required and the *C*-factor is 2000, the maximum flying height (above the terrain) would be 2000 m. One could always be conservative and fly at a lower height, resulting in more photographs at a larger scale. Table 56.2 gives a series of representative parameters for some of the traditional mapping scales used in the U.S. Conventional forward overlap between successive photographs is 60%. Conventional side overlap is 30%. In Table 56.2 the enlargement factor is the factor from negative scale to map scale. Figure 56.16 shows the geometry for determining the base, *B*, or forward gain, from the ground dimension of the photograph coverage, *W*, and the forward overlap as a fraction of *W* (60% is a common value). The following equation gives *B* as described (*OL* = 0.6, the forward overlap fraction):

$$B = (1 - OL) \times W \tag{56.71}$$

The photograph coverage, *W*, is related to the actual photograph dimension, *w*, by the scale
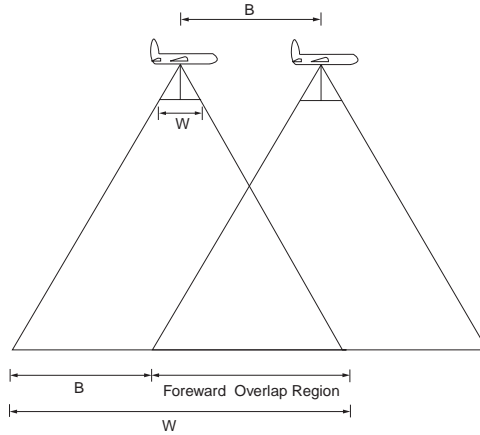
$$W = \text{Scale} \times w \tag{56.72}$$

FIGURE 56.16   Forward overlap geometry.

TABLE 56.2   Typical Flight Parameters (Length Units: Feet)

Map Scale	Ratio	Cont. Intvl.	Neg. Scale	Engl. Factor	H	C-Factor	W	Fwd. Gain	Flt. Line Spc.
1″ = 50′	600	1	4,200	7 ×	2,100	2100	3,150	1,260	2,205
1″ = 100′	1,200	2	7,800	6.5 ×	3,900	1950	5,850	2,340	4,095
1″ = 200′	2,400	5	12,000	5 ×	6,000	1200	9,000	3,600	6,300
1″ = 400′	4,800	5	16,800	3.5 ×	8,400	1680	12,600	5,040	8,820
1″ = 1000′	12,000	10	30,000	2.5 ×	15,000	1500	22,500	9,000	15,750
1″ = 2000′	24,000	10	384,000	1.6 ×	19,200	1920	28,800	11,520	20,160
1″ = 4000′	48,000	20	57,600	1.2 ×	28,800	1440	43,200	17,280	30,240

The photograph dimension, *w*, is approximately 23 cm for standard aerial film. An analogous situation is shown in Fig. 56.17 for determining *S*, the distance between flight lines, from *W* and from the side overlap as a fraction of *W* (30% is a common value). The following equation gives *S* (*SL* = 0.3, the side overlap fraction):

$$S = (1 - SL) \times W \qquad (56.73)$$

Control Points

Horizontal control points (*XY* or *NE*) allow proper scaling of the model data collected from the stereo photographs. Vertical control points (*Z* or *h*) allow the proper terrain slopes to be preserved. Classical surveying techniques determined these two classes of control separately and the distinction is still made. Even with GPS-derived control points, which are inherently 3-D, the elevation or *Z*-dimension is often disregarded because we lack an adequate geoid undulation map to convert ellipsoid heights to orthometric (sea level) referenced heights. Control points may be targeted or signalized before the photo flight so that they are visible in the images, or one may use *photo-ID* or natural points such as manhole covers or sidewalk corners. Coordinates are determined by field survey either before or after the flight, though usually after in the case of photo-ID points. If artificial targets are used, the center panel (at image scale) should be modestly larger than the measuring mark in the stereocomparator or plotter. There should also be two or more prominent legs radiating from the target center panel to allow easy identification in the photographs. Each stereo model requires an absolute minimum of two horizontal control points and three vertical control points. These could all be established by field techniques but this would usually be too expensive. More commonly a sparser network of control points is established, and the needed control in between the field control is determined by bridging or aerial triangulation. When aerial triangulation
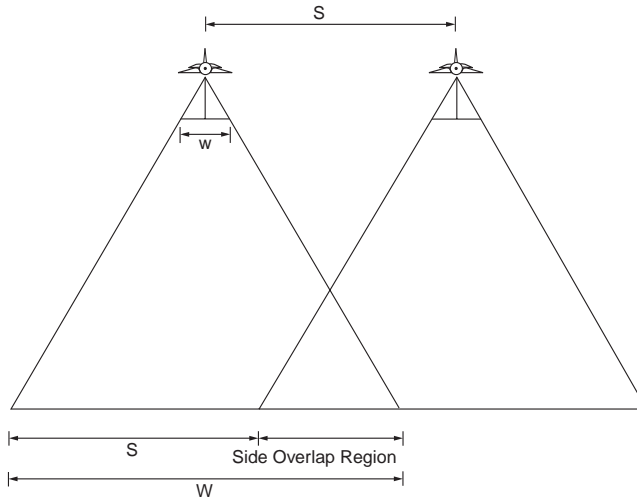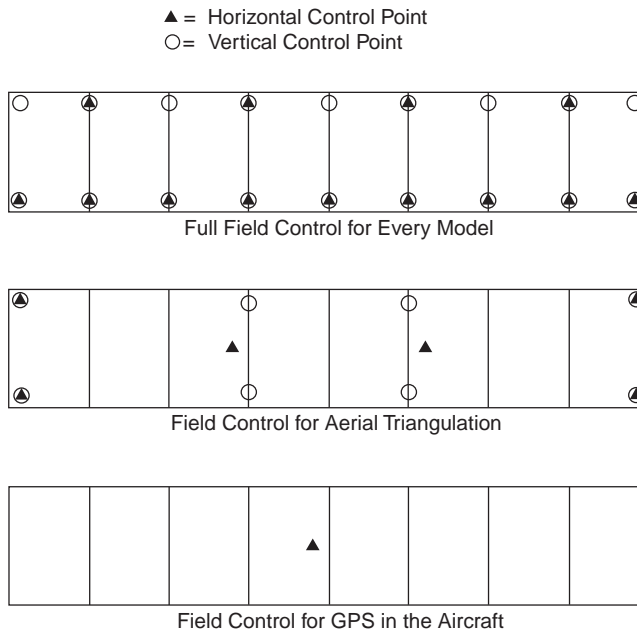
FIGURE 56.17 Side overlap geometry.

FIGURE 56.18 Control point distributions.

is used, artificial marking or pugging of the film diapositives is necessary to create points for the operator to use for absolute orientation. When GPS is used in the aircraft to determine exposure station coordinates independently, then in theory, assuming one has photo strips crossing at near right angles, no ground control is needed. Without crossing strips, at least one control point is necessary. In practice the best number to use will probably fall between 1 and the larger number necessary for ground control-based aerial triangulation. These control point requirements are shown in Fig. 56.18.

56.8 Close-Range Metrology

Close range metrology using photogrammetry offers some unique advantages and presents some equally unique challenges. The advantages are (1) it is a noncontact measurement method, (2) photography can be acquired relatively quickly in dangerous or contaminated areas, (3) the photographs represent an archival record in case any dimensions are disputed, and (4) the camera can be brought to the object, rather than the other way around (as with a coordinate measuring machine). Accuracies obtainable using close-range photogrammetry (as in aerial photogrammetry) would fall in the range of 1/5,000 to 1/100,000 of the object distance.

Equipment

Cameras are typically small format, usually with 70-mm film size. Stereometric cameras are mounted rigidly in pairs, so that fewer parameters are required in solving for the orientations. The cameras used would preferably be metric, although with very special handling, and much extra effort, nonmetric cameras can be used for some applications. There should be at least some sort of fiducial marks in the focal plane, preferably a reseau grid covering the image area. Film flatness can be a severe problem in nonmetric cameras. Ideally there should be a vacuum back permitting the film to be flattened during exposure. There should be detents in the focus settings to prevent accidental movement or slippage in the setting. Calibration will be carried out at each of these focus settings, resulting in a set of principal distances and lens distortion curves for each setting. These are often computed using added parameters in the bundle adjustment, performed by the user, rather than sending the camera to a testing laboratory. Lighting can be a problem in close-range photogrammetry. For some objects and surfaces which have very little detail and texture, some sort of "structured lighting" is used to create a texture. To obtain strong geometry for the best coordinate accuracy, highly convergent photography is often used. This may prevent conventional stereo viewing, leading to all photo observations being made in monoscopic view. This introduces an additional requirement for good targeting, and can make advantageous the use of feature equations rather than strictly point equations.

Applications

Close-range photogrammetry has been successfully used for tasks such as mapping of complex piping systems, shape determination for parabolic antennas, mating verification for ship hull sections, architectural/restoration work, accident reconstruction, and numerous medical/dental applications.

56.9 Remote Sensing

Remote sensing is considered here in its broad sense, including the photogrammetric aspects of using nonframe imagery from spaceborne or airborne platforms. A thorough treatment must include the metric aspects of the image geometry and the interpretive and statistical aspects of the data available from these sources.

Data Sources

Following is a partial listing of sensors, with associated platform and image descriptions. These sensors provide imagery which could be used to support projects in civil engineering.

1. MSS, multispectral scanner, Landsat 1-5, altitude 920 km, rotating mirror, telescope focal length 0.826 m, IFOV (instantaneous field of view) 83 m on the ground, gray levels 64, image width 2700 pixels, 4 spectral bands: 0.4–1.0 micrometers
2. TM, thematic mapper, Landsat 4-5, altitude 705 km, rotating mirror, telescope focal length 1.22 m, IFOV 30 m on the ground, gray levels 256, image width 6000 pixels, 6 spectral bands: 0.4–0.9, 1.5–1.7, 2.1–2.4, 10.4–12.5 micrometers
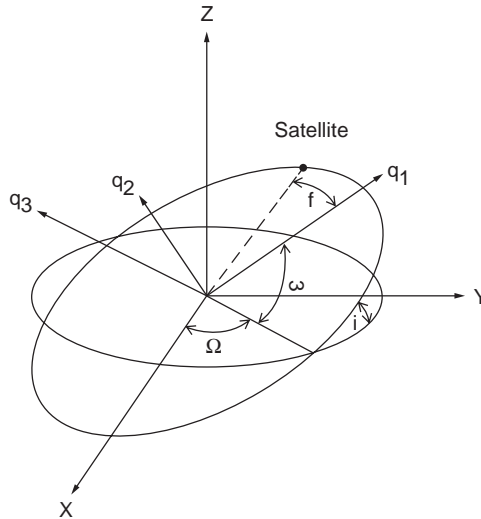
FIGURE 56.19   Orbital parameters.

3. SPOT, SPOT 1-3, panchromatic (multispectral not described), altitude 822 km, pushbroom, telescope focal length 1.082, IFOV 10 m on the ground, gray levels 256, image width 6000 pixels, 1 panchromatic band 0.55–0.75 micrometers plus 3 multispectral bands, off-nadir viewing capability for stereo

Geometric Modeling

Coordinate Systems

The two primary coordinate systems needed for describing elliptical orbits as occupied by imaging sensor platforms are the *XYZ* system and the $q_1q_2q_3$ system shown in Fig. 56.19. The *XYZ* system has the *XY* plane defined by the earth equatorial plane, *Z* oriented with the spin axis of the earth, and with *X* through the vernal equinox or the first point of Aries. The second system has q_1q_2 in the orbit plane with origin at one focus of the ellipse, as shown in Fig. 56.19. q_1 is along the semimajor axis; q_2q_3 are as described above to define a right-handed coordinate system.

Orbital Mechanics

The gravitational attraction between the earth and an orbiting satellite is

$$F = \frac{GMm}{r^2} \tag{56.74}$$

where m = the mass of the satellite
          r = the geocentric distance to the satellite
          GM = the earth's gravitational constant:

$$GM = 3,986,004 \times 10^8 \, \text{m}^3\text{s}^{-2}$$

For a small mass satellite,

$$\mu = GMm \approx GM \tag{56.75}$$

Expressing the position vector in 3-D coordinates and the position in the orbital plane in polar coordinates, extracting an expression for the acceleration, and further manipulating the expressions yields a differential equation whose solution is a general conic section in polar coordinates:

$$r = \frac{p}{1 + e \cos(f)} \tag{56.76}$$

where f is the true anomaly as shown in Fig. 56.19, p is the semilatus rectum, and e is the eccentricity. For $e > 1$ the orbit is a hyperbola, for $e = 1$ the orbit is a parabola, and for $e < 1$ the orbit is an ellipse. The use of the term *anomaly* as a synonym for angle is a vestige of Ptolemaic misconceptions, originally implying an angle which did not vary linearly with time. The true anomaly, f, the eccentric anomaly, E, and the mean anomaly, M, are related by Kepler's equation:

$$M = E - e \sin E \tag{56.77}$$

and

$$\tan\left(\frac{f}{2}\right) = \sqrt{\frac{1 + e}{1 - e}} \tan\left(\frac{E}{2}\right) \tag{56.78}$$

The point on an elliptical orbit furthest from the earth is the *apogee*; the point closest is the *perigee*. At perigee the true anomaly is zero; at apogee it is 180 degrees. An orbit is specified when the satellite position and velocity vectors, $\mathbf{r}$ and $\dot{\mathbf{r}}$ are given at a particular time. A more common set of six parameters to specify an orbit are the classical Keplerian elements $\{a, e, T_0, \Omega, i, \text{ and } \omega\}$, where a is the semimajor axis, e is the eccentricity, T_0 is the time of perigee passage, Ω is the right ascension of the ascending node (shown in Fig. 56.19), i is the inclination, and ω is the argument of the perigee.

In order to obtain the Kepler elements from the position and velocity vectors at a given time, the following equations can be used. The angular momentum vector is

$$\mathbf{h} = \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} = \mathbf{r} \times \dot{\mathbf{r}} \tag{56.79}$$

This yields

$$\Omega = \tan^{-1}\left(\frac{h_x}{-h_y}\right) \tag{56.80}$$

$$i = \tan^{-1}\left(\frac{\sqrt{h_x^2 + h_y^2}}{h_z}\right) \tag{56.81}$$

With v representing the velocity,

$$a = \frac{r}{2 - (rv^2/\mu)} \tag{56.82}$$

$$e = \sqrt{1 - \frac{h^2}{\mu a}} \tag{56.83}$$

Obtaining the eccentric anomaly from

$$\cos E = \frac{a - r}{ae} \tag{56.84}$$

the true anomaly is found from

$$f = \tan^{-1}\left(\frac{\sqrt{1 - e^2}\sin E}{\cos E - e}\right) \tag{56.85}$$

Defining an intermediate coordinate system in the orbit plane such that p_1 is along the nodal line,

$$p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = R_1(i)R_3(\Omega)r \tag{56.86}$$

$$\omega + f = \tan^{-1}\left(\frac{p_2}{p_1}\right) \tag{56.87}$$

T_0 can be obtained from

$$M = \sqrt{\frac{\mu}{a^3}}(t - T_0) \tag{56.88}$$

To go in the other direction, from the Kepler elements to the position and velocity vectors, begin with Eq. (56.88) to obtain the mean anomaly, then solve Kepler's equation, Eq. (56.77), numerically for the eccentric anomaly, E. The true anomaly, f, is found by Eq. (56.78). The magnitude of the position vector in the orbit plane is

$$r = \frac{a(1 - e^2)}{1 + e\cos(f)} \tag{56.89}$$

the position vector is

$$r = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} r\cos(f) \\ r\sin(f) \\ 0 \end{pmatrix} \tag{56.90}$$

and the velocity vector is

$$v = \sqrt{\frac{\mu}{p}}\begin{pmatrix} -\sin(f) \\ e + \cos(f) \\ 0 \end{pmatrix} \tag{56.91}$$

where p, the semilatus rectum, is given by

$$p = a(1 - e^2) \tag{56.92}$$

The rotation matrix relating the orbit plane system and the vernal equinox, spin-axis system is given by

$$r_{XYZ} = Rr_{q_1q_2q_3} \tag{56.93}$$

where R is

$$R = R_3(-\Omega)R_1(-i)R_3(-\omega) \tag{56.94}$$

This R can be used to transform both the position and velocity vectors above into the XYZ "right ascension" coordinate system defined above.

Platform and Sensor Modeling

If we consider a linear sensor on an orbiting platform, with the sensor oriented such that the long dimension is perpendicular to the direction of motion (i.e., pushbroom), we can construct the imaging equations in the following way. Each line of imagery, corresponding to one integration period of the linear sensor, can be considered a separate perspective image, having only one dimension. Each line would have its own perspective center and orientation parameters. Thus, what may appear to be a static image frame is, in fact, a mosaic of many tiny "framelets." These time dependencies within the image "frame" are the result of normal platform translational velocity, angular velocity, and additional small, unpredictable velocity components. The instantaneous position and orientation of a platform coordinate system is shown in Fig. 56.20.

The equations which relate an object point, $(X_m, Y_m, Z_m)^t$, and the time-varying perspective center, $(X_s(t), Y_s(t), Z_s(t))^t$ are of the form

$$\begin{bmatrix} 0 \\ y \\ -f \end{bmatrix} = k \times R_{III} \times R_{II} \times R_I \times \begin{bmatrix} X_m - X_s(t) \\ Y_m - Y_s(t) \\ Z_m - Z_s(t) \end{bmatrix} \tag{56.95}$$

R_{III} is a platform to sensor rotation matrix which should be fixed during a normal frame. This could accommodate off-nadir view angle options, for instance. R_{II} transforms the ideal platform to the actual
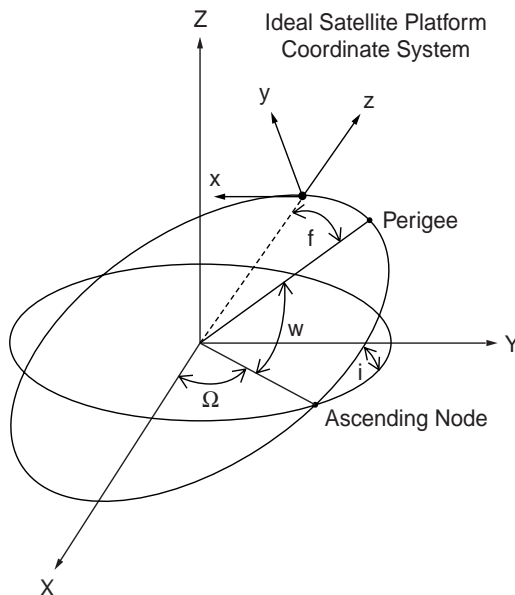

FIGURE 56.20 Imaging satellite orientation.

platform and the time dependency in its angles would probably take the form of low degree polynomial functions. R_I represents the time-dependent rotation from a space-fixed geocentric coordinate system to the instantaneous platform coordinate system. In a manner similar to the previous section, this matrix would be the composition of the rotations,

$$R_I = R_z(\pi) \times R_y\left(\frac{\pi}{2} - \omega - f(t)\right) \times R_x\left(i - \frac{\pi}{2}\right) \times R_z(\Omega) \qquad (56.96)$$

with the orbit elements as described in the previous section. The Cartesian coordinates on the right side of Eq. (56.95) are with respect to a space-fixed coordinate system (i.e., not rotating with the earth). In order to transform a point in a terrestrial, earth-fixed coordinate system, a coordinate rotation, R_0, is necessary to account for the earth's rotation, polar motion, precession, and nutation. For $(X_e, Y_e, Z_e)^t$ in an earth-fixed system,

$$\begin{bmatrix} X_m \\ Y_m \\ Z_m \end{bmatrix} = R_0 \times \begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix} \qquad (56.97)$$

Analogous to the decomposition of conventional orientation in photogrammetry into exterior and interior components, we could construct here a decomposition of the imaging Eq. (56.95) into a platform component and a sensor component. Eliminating the nuisance scale parameter in Eq. (56.95) would yield two equations per object point, per image "frame." Unknown parameters would consist of a user-selected subset of the platform and sensor orientation parameters, some time varying, the orbit parameters, and the object point parameters. These equation systems could be formed for a single image (resection), a set of sequential images along the orbit path (strip), or an arbitrary collection of adjacent frames (block).

The imminent arrival of commercially available satellite data in the 1- to 3-meter pixel range means that this kind of image modeling will become even more important in the future as digital sensors continue to slowly supplant film-based photography.

Interpretive Remote Sensing

Remote sensing, as the term is usually employed, implies the study of images to identify the objects and features in them, rather than merely to determine their size or position. In this way, remote sensing is a natural successor to the activity described by the slightly outdated and more restrictive term *photo interpretation*. Remote sensing almost always implies digital data sources and processing techniques. It often implies multiple sensors in distinct spectral bands, all simultaneously imaging the same field of view. Along with the multispectral concept, there is no restriction that sensors respond only within the range of visible radiation. Indeed, many remote sensing tasks rely on the availability of a wide range of spectral data. Witness the recent move toward "hyperspectral" data, which may have hundreds of distinct, though narrow, spectral bands.

Systems have been deployed with sensitivities in the ultraviolet, visible, near infrared, thermal infrared, and microwave frequencies. Microwave systems are unique in that they can be *active* (providing source radiation as well as detection), whereas all of the others listed are strictly *passive* (detection only). Active microwave systems include both SLAR, side-looking airborne radar, and synthetic aperture radar.

Multispectral Analysis

Multispectral image data can be thought of as a series of coregistered image planes, each representing the scene reflectance in a discrete waveband. At each pixel location, the values from the spectral bands constitute an *n*-dimensional data vector for that location. Consider an example with two spectral bands in which the scene has been classified a priori into three categories: water (W), agricultural land (A),
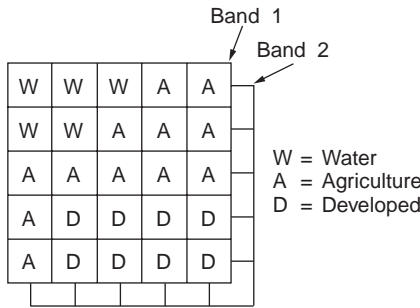
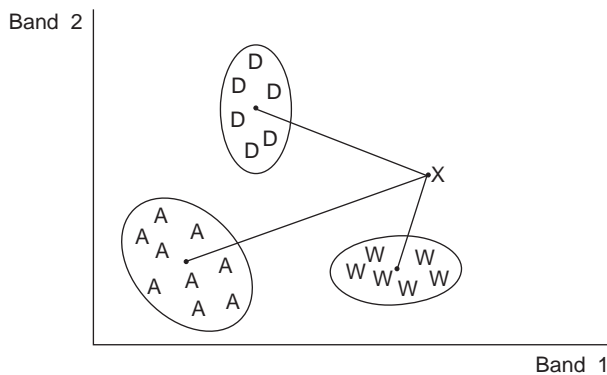**FIGURE 56.21** Training sample for supervised classification.

FIGURE 56.22 Clustering of 2 band data in scatter diagram.

and developed land (*D*), as shown in Fig. 56.21. From this *training sample* one can (assuming normality) construct a mean vector and a variance/covariance matrix for each class. New observation vectors could then be assigned to one of the defined classes by selecting the minimum of the Mahalanobis function,

$$D^2 = (X - \mu_i)^t \Sigma_i^{-1} (X - \mu_i) + \ln|\Sigma_i| \tag{56.98}$$

where μ_i and Σ_i are evaluated for each category. Clustering of feature classes is shown in Fig. 56.22. This is an example of a maximum likelihood classifier, using *supervised* classification. Other techniques can be used such as discriminant functions and Bayesian classification. In addition to statistical and heuristic approaches to pattern recognition, one can also employ syntactic methods, wherein a feature is identified by its context or relationship to adjacent or surrounding features.

Change Detection

In order to detect changes on a portion of the earth's surface it is necessary to have imagery at different times. In addition, when possible, it is desirable to minimize apparent, but spurious, differences due to season, view altitude, and time of day. In order to make a proper pixel-by-pixel comparison, both images should be resampled and brought to the same map projection surface. This is best done by using identifiable ground control points (road intersections, etc.) and a modeling of the sensor and platform positions and orientations. Less effective registration is sometimes done using two-dimensional polynomial functions and a "rubber sheet" transformation to the control points. Once the scenes have been brought to a common coordinate base, a classification is made individually on each scene. A pixel-by-pixel comparison is made on the classification results, and a "change" layer can be generated where differences are encountered.

Microwave Remote Sensing

Microwave radiation has the useful characteristic of penetrating clouds and other weather conditions which are opaque to visible wavelengths. Aircraft-based imaging radars are thus able to acquire imagery under less restrictive flight conditions than, for example, conventional photography. The geometry of the radar image is fundamentally different from the near perspective geometry of other sensor systems. In both real aperture side-looking airborne radar and synthetic aperture radar, imagery is presented as a succession of scan lines perpendicular to the flight direction. The features displayed along the scan lines have positions proportional to either the slant range or the ground range, and the image density or gray level is related to the strength of the return. Radar imagery can be acquired of the same area from two flight trajectories, inducing height-related parallax into the resulting stereo pair. With the proper imaging equations, stereo restitution and feature compilation can be carried out, giving rise to the term *radargrammetry*.

Until recently, remote sensing activities were the exclusive domain of a few highly industrialized countries. Now a number of countries and even some commercial ventures are becoming involved in systems to provide remote sensing imagery. For the civil engineering community, this increased availability of data can be very beneficial.

References

Burnside, C.D. 1985. *Mapping from Aerial Photographs*. John Wiley & Sons, New York.

Colwell, R.N., ed. 1983. *Manual of Remote Sensing*. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Ebner, H., Fritsch, D., and Heipke, C. 1991. *Digital Photogrammetric Systems*. Wichman, Karlsruhe.

Faugeras, O. 1993. *Three-Dimensional Computer Vision*. MIT Press, Cambridge, MA.

Karara, H.M., ed. 1989. *Non-Topographic Photogrammetry,* 2nd ed. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Kraus, K. 1993. *Photogrammetry*. Dummler Verlag, Bonn.

Leick, A. 1990. *GPS Satellite Surveying*. John Wiley & Sons, New York.

Moffitt, F.H., and Mikhail, E.M. 1980. *Photogrammetry,* 3rd ed. Harper & Row, New York.

Pease, C.B. 1991. *Satellite Imaging Instruments*. Ellis Horwood, Chichester.

Richards, J.A. 1993. *Remote Sensing Digital Image Analysis: An Introduction,* 2nd ed. Springer-Verlag, New York.

Slama, C.C., ed. 1980. *Manual of Photogrammetry,* 4th ed. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Wolf, P.R. 1983. *Elements of Photogrammetry*. McGraw-Hill, New York.

Further Information

The following journals are useful sources of information:

Photogrammetric Engineering and Remote Sensing, American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

The Photogrammetric Record, The Photogrammetric Society, London.

ISPRS Journal of Photogrammetry and Remote Sensing (formerly *Photogrammetria*), Elsevier Science Publishers B.V., Amsterdam, The Netherlands.

CISM Journal (formerly *Canadian Surveyor*), Canadian Institute of Surveying and Mapping, Ottawa, Canada.

Journal of Surveying Engineering, American Society of Civil Engineers, New York.

The following organizations provide valuable technical and reference data:

American Society of Photogrammetry and Remote Sensing, 5410 Grosvenor Lane, Suite 210, Bethesda, MD 20814

American Congress on Surveying and Mapping, 5410 Grosvenor Lane, Bethesda, MD 20814

AM/FM International, 14456 E. Evans Ave., Aurora, CO 80014
U.S. Geological Survey, EROS Data Center, Sioux Falls, SD 57198
U.S. Geological Survey, Earth Science Information Center, 53. National Center, Reston, VA 22092
SPIE, The International Society for Optical Engineering, P.O. Box 10, Bellingham, WA 98227
American Society of Civil Engineers, 345 East 47th Street, New York, NY