Three-Dimensional Reconstruction from Two-Point Perspective Imagery

James R. Williamson and *Michael H. Brill*
Science Applications International Corporation, Falls Church, VA 22046

ABSTRACT: Although three-dimensional object-space reconstuction is generally not done from a single image, there are special cases in which *a priori* information (or reasonable assumptions) about the viewed objects allows such a reconstruction. In particular, man-made objects tend to have recognizable geometry, e.g., planar surfaces, straight-line edges, and right angles. These features can be used in a three-dimensional reconstruction, which is then complete except for a scale factor. If in addition one dimension of the object is known, the scale factor is also recoverable. The present article discusses useful graphical techniques for reconstructing three-dimensional information from a single image, using for illustration a two-point perspective image of a rectangular solid. Projective-geometric justification is presented for the constructions, and it is suggested that the techniques are applicable to forensic photogrammetry.

INTRODUCTION

PERSPECTIVE is the basis for photogrammetric analysis of close-range imagery (Slama, 1980). Much useful information can be obtained from close-range perspective imagery of three-dimensional man-made objects. It is often feasible to obtain from such images the dimensions of objects that are difficult or impractical to measure physically. In close-range forensic photogrammetry, there is usually very little or no information provided concerning the camera orientation relative to the object imaged. In addition, many close-range forensic images may have no known interior orientation, often being cropped enlargements or taken with cameras whose parameters are inaccessible to the photogrammetrist (Beamish, 1984). Finally, such imagery is rarely provided as stereo pairs. In compensation for the lack of parametric values, there are usually recognizable features that allow a three-dimensional reconstruction from a single photograph. The single-image perspective photogrammetric analysis is an ideal approach to use for limited close-range (forensic) imagery because it uses the geometrical properties of the objects imaged, particularly objects composed of rectangular solids. The perspective geometry of such objects allows determination of the necessary interior and exterior parameters using a few known dimensions and some major assumptions. Some references (e.g., McCartney, 1963; Gracie *et al.*, 1967; Kelley, 1978-1983, Busby, 1981; and Novak, 1986) provide methods to exploit three-point perspective, i.e., in which none of the object's three major orthogonal planes are parallel to the image plane, and hence there are three vanishing points. A few references (e.g., Walters and Bromham, 1970; Busby, 1981; and Kelley, 1978-1983) provide useful techniques for exploiting two-point perspective, in which one axis of the rectangular solid is parellel to the image plane and hence there are two vanishing points. Ethrog (1984) has developed *analytical* methods for determining camera attitude, focal length, and principal point values from parallel and perpendicular lines. Of course, these methods can be applied to the extraction of three-dimensional information from a single image, whether it has two-point or three-point perspective. In the present article, we discuss some *graphical* methods useful in reconstructing three-dimensional geometry from two-point perspective images.

We bring our graphical methods to bear on the three-dimensional reconstruction of a rectangular solid determining the object-space coordinate planes. For such a solid appearing in two-point perspective, we will show that *a priori* knowledge of a single diagonal angle is enough to reconstruct a scale model of the solid. The coordinates of two object-space points or one object-space dimension will be required to provide scale. In the process of constructing the object-scale model, we will determine vanishing points, the camera station (in the image and in object space), the perspective principal point, the perspective focal length, and the camera rotation angles of azimuth, tilt, and swing. By using the lens center as a projection point, we reconstruct three true orthographic views of the object on paper. In contrast with other methods (Gracie *et al.*, 1967; Busby, 1981; Novak, 1986), this reconstruction will *not* require a full-format image. Our methods are adopted from the geometrical procedures used to produce architectural drawings (McCartney, 1963; Walters and Bromham, 1970), but, of course, our goal is to infer the three-dimensional geometry from an image rather than to create the image from three-dimensional information.

Three-dimensional reconstruction from an image requires knowledge of the geometry of perspective (Duda and Hart, 1973) whereby lines that are parallel in object space appear on an image as lines that converge to a single point, a vanishing point. In a three-dimensional cartesian coordinate system, there is one vanishing point for each coordinate axis: all lines parallel to an axis will converge at the vanishing point for that axis. According to the juxtaposition of the camera and object, it is possible to have one-, two-, or three-point perspective imagery (see Figure 1). Of course, in general there are as many vanishing points for an image as there are sets of parallel lines, but the rectangular solid distinguishes three of these sets. In the present article, we discuss the analysis of images that have two-point perspective.

THE FUNDAMENTALS OF GRAPHICAL TWO-POINT PERSPECTIVE

The most basic construction in perspective geometry is a group of imaged parallel lines converging at some point referred to as a vanishing point (see Figure 1.). From this construction evolves the interaction of vanishing points, the principal point, the focal length, and the camera station of the perspective image-space model (Gracie *et al.*, 1967; Kelley, 1978; Slama, 1980; Moffitt and Mikhail, 1980).

To grasp the concept of two-point perspective, it is important to compare the geometry of the graphical image-space models of two- and three-point perspective. The edge views for two-point and three-point perspective image-space models are shown in Figure 2a and 2b. When the tilt (t) is 90° (Figure 2b) the image plane becomes parallel to the Nadir line (the line CS-VPZ); the point VPZ goes to infinity; and the points VPX, VPY, and PP become coincident with THL (true horizon line). Note that, by definition, THL is the trace of the horizontal plane, through the camera station, across the image plane. As shown in Figure 2c, the line between VPX and VPY (labeled THL) is the diameter of a circle and the CS is a point on the circle. To illustrate this graphically while keeping the relationships of the geometry true, the semicircle from the VPX to the VPY through the CS is rotated 90° about the THL until the semicircle lies in the image plane. True relationships that were constructed in Figure 2b (the side view) may now be shown in Figure 2c (the front view). In

0099-1112/87/5303–331$02.25/0

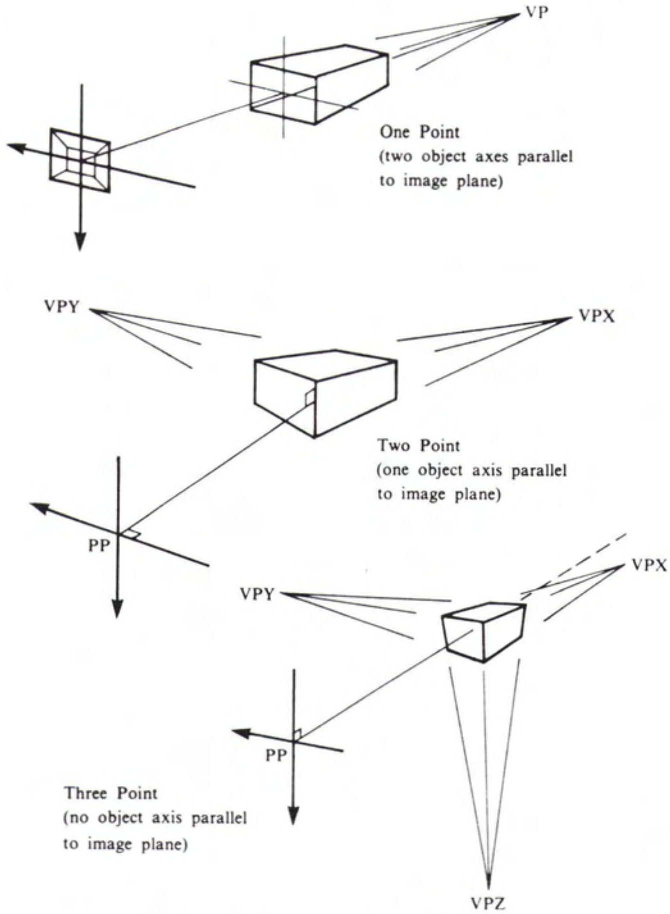
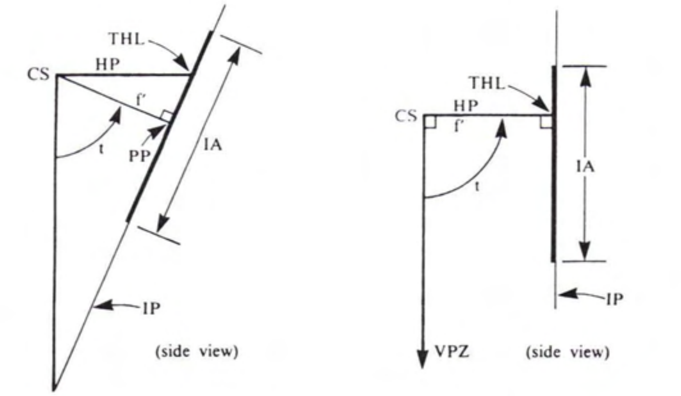
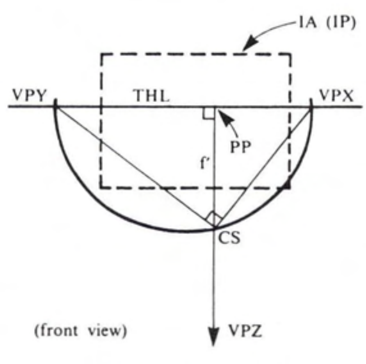


FIG. 1. One-, two-, and three-point perspective.



2a. Three-point perspective edge view

2b. Two-point perspective edge view



2c. Two-point perspective horizontal plane in image plane

Definitions	
CS	Camera Station (lens effective nodal point)
HP	Horizontal Plane
IA	Image Area (not necessarily full format)
THL	True Horizon Line
PP	Principal Point

FIG. 2. Graphical relationships of image space model.

fact, any image-model vertical relationships that can be constructed between CS and the image-space model THL may still be shown in the rotated horizontal plane of Figure 2c, as if the horizontal plane had not been rotated. This is a standard procedure of descriptive geometry and is very practical in graphical photogrammetric solutions.

In our two-point perspective analysis (see Figure 3), parallel planes, such as ABCD and EFGH (top and bottom of the rectangular solid) are defined to be horizontal. The edges of the base — HG and HE — define the X and Y axis directions, and point to the associated vanishing points VPX and VPY of the image. Between the vanishing points is the THL. Because the image plane is perpendicular to the object horizontal planes, all vertical lines of the object are parallel to the image plane, and the principal ray points directly at the THL — defining a tilt of 90°. (Actually, two-point graphical procedures are usable whenever the tilt is 90° + or − 3°.) Thus, the vertical imaged lines have no graphically definable vanishing point, because such a point would be an infinite distance above or below the THL.

If a two-point perspective image has full format, its perspective principal point (PP) may best be estimated as the image's center of format (center of vision, see Busby (1981); image-center, see Novak (1986)). However, even without full format, a two-point perspective image allows three-dimensional reconstruction given some other *a priori* knowledge. The reconstruction starts by determining the location of the camera station (CS), and thereby the PP and effective focal length (f'). The CS can be found by locating the vanishing point of horizontal- or vertical-plane diagonals, given a known angle (θ_h or θ_v) of the diagonal in object space (see Figures 4 and 5). The easiest diagonal to use is the diagonal of a horizontal or vertical square enclosing a circle. A man-hole cover or an automobile wheel rim are such circles and can graphically (using VPX and VPY) be enclosed in a perspective square whose diagonal angle is known.

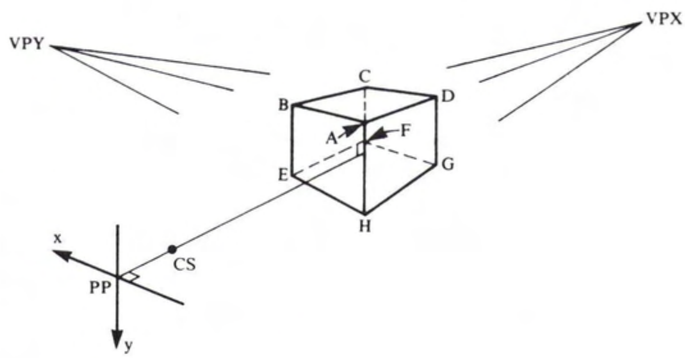


FIG. 3. Two-point perspective and vertical line illustration.

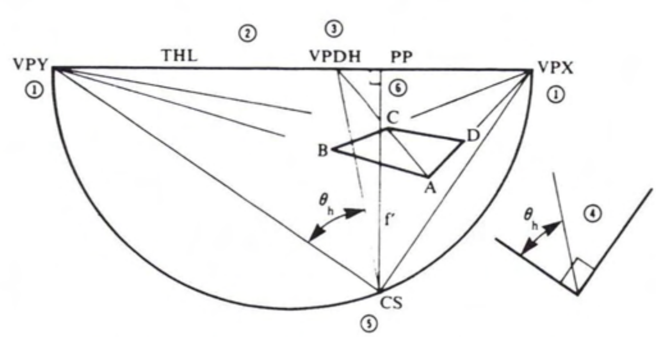


FIG. 4. Six steps using horizontal-plane diagonals.

A procedure for using a horizontal-plane diagonal to locate

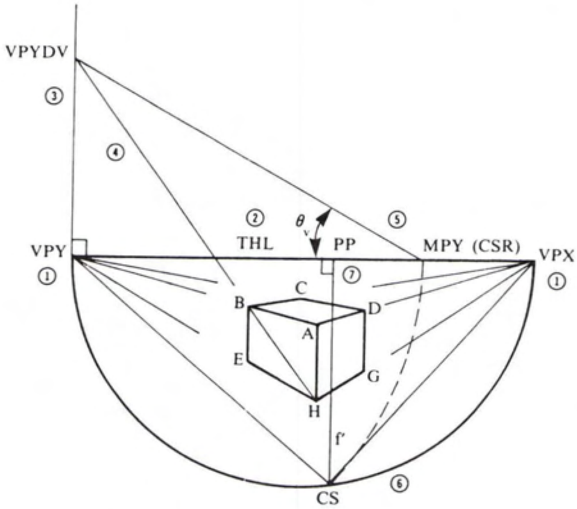


FIG. 5. Seven steps using vertical-plane diagonals.

the CS and PP is illustrated in Figure 4. There are six steps in this procedure. Using any horizontal plane (with known diagonal angle) on the object, such as ABCD in Figure 4, the first step is to locate the vanishing points VPX and VPY for the X and Y coordinate axes. The second step is to draw the THL between the two vanishing points and semicircle with the THL as the diameter. With two-point perspective it is best to draw the semicircle beneath the THL, although the solution can be worked with the semicircle above the THL. Third, locate the intersection point of the THL with the horizontal-plane diagonal line (AC) extended. This is the vanishing point for the horizontal-plane diagonal (VPDH). Fourth, on a clear plastic overlay scribe two lines at right angles and between them the known-angle line (θ_h = 45° for our square). The graphical intersection of these three lines (each passing through a vanishing point) on the semicircle will be the CS. Each line must be long enough to extend beyond its respective vanishing point, allowing easy traverse of the semicircle. Fifth, move the position of the CS around the semicircle, and, at the instant when each line simultaneously passes through a vanishing point, you have located the perspective image position of the CS. Sixth, draw a line perpendicular to the THL and through CS. The intersection of the line with the THL locates the PP and the f'. This completes the six-step procedure.

The construction shown in Figure 4 can be understood geometrically by rotating the semicircle (with CS) about the axis THL, 90° out of the paper (image plane). CS now represents the true-ratio (three-dimensional) camera station location and PP is the true principal point as drawn. The line CS-VPY is parallel to the true line AB for the horizontal rectangle, and CS-VPX is parallel to the true line AD. The line CS-VPDH, at angle θ_h from the line CS-VPY, is parallel to the true diagonal AC, hence the vanishing-point interpretation of VPDH.

Equally straightforward (see Figure 5) is the use of a vertical-plane diagonal to locate the CS and PP. There are seven steps in this procedure. The first two steps are the same as the steps for horizontal diagonals: locate the X and Y vanishing points (VPX and VPY) and draw the semicircle below the THL. The third step is to draw a perpendicular to the true horizon from either the X or Y vanishing point, depending upon whether the vertical diagonal is in the YZ or XZ plane. In Figure 5, the extended diagonal (HB) is in the YZ plane and passes over the Y vanishing point, so the perpendicular is through VPY. Fourth, extend the diagonal to intersect with the perpendicular just constructed. The intersection is the vanishing point for the vertical diagonal (VPYDV). Fifth, find the position on the THL such that lines to VPYDV and VPY subtend the known angle θ_v relative to this point (in the case of Figure 5, θ_v is 30°). The point so located is normally referred to as a Y measuring point (MPY), but it is also the camera station revolved (CSR). The point VPY is the center of revolution for constructing the CS from the CSR. Therefore, sixth, draw an arc from MPY that intersects with the semicircle. The intersection point is the CS. Step seven is to draw a perpendicular from the THL through CS, thus locating PP on the THL. This completes the seven-step procedure.

To support accuracy for the solution, all available parallel diagonal lines should be used to locate the horizontal- or vertical-plane diagonal vanishing point. If possible, at least three lines should be used. In fact, if it were possible to use both a horizontal- and vertical-plane diagonal vanishing point to locate the camera station, it would establish a greater confidence level in the solution than if just one diagonal vanishing point were used.

The construction of Figure 5 can be understood geometrically (after the fact) by rotating the semicircle (with CS) about THL, 90 degrees out of the paper (image plane). As with Figure 4, CS and PP are thereby rendered in accurate image-model three-dimensional positions. Because the diagonal HB is in the same plane as HE and AB, and this plane is vertical, it follows that the vanishing point for HB is in the image-vertical direction from VPY (the vanishing point of HE and AB). Hence, the vanishing point for HB is found by extending HB to intersect (at VPYDV) the perpendicular line through VPY, as illustrated. With the camera station in its true image-model three-dimensional position (after the rotation described at the beginning of this paragraph), the line from true-CS to VPY and the line from true-CS to VPYDV make the same angle θ_v as the lines HE and HB on the object. This is because a vanishing point is a meeting of the imaged parallel lines of object space; hence, true CS-VPY is parallel to true HE, and true CS-VPYDV is parallel to true HB. To depict these lines graphically (from true CS with the angle θ_v), rotate the true CS about VPY-VPYDV until these lines lie in the plane of THL and VPYDV (the image plane). The point, on the THL, now occupied by CS is the camera station revolved (CSR or measuring point Y — MPY). This location of CSR is exactly the same as obtained by rotating the CS in the image plane about VPY to the THL (as shown in Figure 5). The converse is true when the location of MPY on the THL is known (having been located by knowing VPYDV and θ_v). The image-plane CS is located by rotating MPY about VPY, until CS is generated at the intersection with the semicircle. A perpendicular from the THL through this newly located CS determines the PP and the f'.

Once the construction of PP is complete, the f' is then measurable as the distance between PP and CS. Also, of the three rotation angles (azimuth, tilt, and swing, as defined by Slama (1980)), the azimuth (a) is the angle between the THL and the line from the VPX to CS. All of these parameters are identified in Figure 6.

This leaves the swing angle (s) to be determined. The swing angle defines the rotation of the image frame relative to the true horizon line. By definition (Slama, 1980), swing is "the angle at the principal point of a photograph which is measured clockwise from the positive y-axis to the principal line at the nadir point." The image coordinate system is defined with its x-axis parallel to the bottom format edge, with its y-axis perpendicular to the x-axis, and with its origin at the principal point. If the camera were level to the horizontal plane of the object-space coordinate system at the time of exposure, the swing angle would be 180 degrees. If the camera were not level to the horizontal plane and the image had dull format, then the swing could easily be measured. If there is no image full format, swing (shown in Figure 6 for definition and clarity) is always taken to be 180°, this choice having no consequences in the graphical solution.

Given CS, VPX, and VPY, it is possible to draw a true-ratio plan view and elevation views of the imaged object. The plan-view construction is shown in Figure 7. Draw perpendiculars to the THL from the object as shown. Using point A' (which is

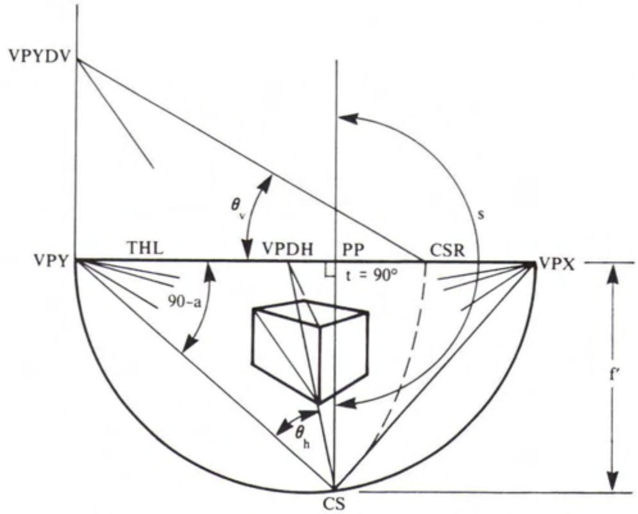


FIG. 6. Two-point perspective parameters identified.

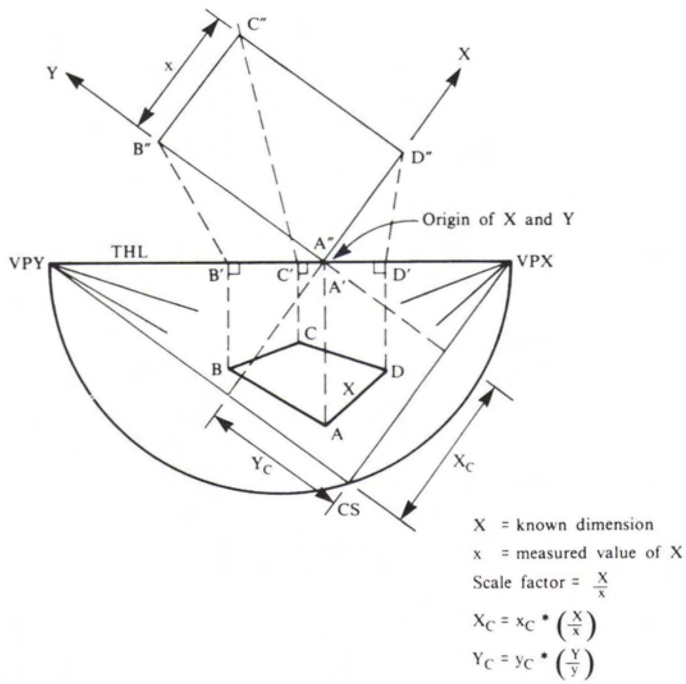


FIG. 7. Two-point perspective plan view and CS coordinates X_C and Y_C constructed to the same scale.

the same as A" because the vertical reference plane is THL in this example) as one corner of the plan view, draw lines parallel to CS-VPX and CS-VPY through point A'. Extend lines from CS through D' and B' until they intersect the newly drawn lines at D" and B", respectively. Lines are drawn perpendicular to A"D" and A"B" through the points D" and B" to intersect on the extended line CS-C'. This forms the corner C". Graphically the accuracy of the construction is visually checked at point C" by how closely the intersections of the three lines cluster. Adjustments may be made to the graphical solution until the intersection at point C" is acceptable.

The construction in Figure 7 is obtained by rotating the plane of the image semicircle into the plane of the plan view, projecting vertical lines to the THL (or plane hinge) and drawing lines from the projection center (CS) through the points on the THL. (Note that the THL is also an edge view of the image plane.) These extended lines intersect the XZ and YZ planes and locate the position of the object vertical lines as points in the plan view. The angles of A"B" and A"D" with respect to the THL are clearly the same as the angles between the THL and the CS-VPY and CS-VPX lines, respectively. In this scaled plan view, with the origin of the object-space coordinate system at point A", the true camera station coordinates, X_c and Y_c, are easily found. Note that the object-space plan view is arbitrarily placed with respect to the THL. The relation of the plan-view scale to the image scale is dictated by the positioning of an additional reference plane parallel to the image plane. In this example the image plane was used as the reference plane.

The construction of a true-ratio side view of a vertical plane of the rectangular solid (see Figure 8) is even easier, and also gives rise to the camera station coordinate Z_c with the same scaling factor. Given VPY, VPX, and CS, extend line HA to meet THL at A'. Extend diagonal HB until it meets the vertical line at VPY. The intersection point is VPYDV. Rotate CS about VPY in the image until the intersection MPY (H") is located on THL. Then, the true length-to-width ratio of the vertical rectangle is generated by H", E", and B". The coordinate Z_c also appears implicitly in Figure 8, for HA'/HA on the image is equal to Z_c/(true HA) in object space. Hence, HA" represents Z_c to the scale of HA. An elevation view of HEBA is obtained by completing the rectangle H"E"B"A" as shown. The construction can be understood in the same way as that of Figure 5, and can be used to construct the other elevation view of the rectangular solid by using VPX instead of VPY.

Although the two side views and the plan view obtained by these constructions are not to the same scale, they offer enough redundancy to convert them to the same scale (e.g., the scale for which AD = 1 cm). Of course, the true length AD is still undetermined by this single-image construction, unless a dimension (say, AD = X in Figure 7) is known. In that case, all the true distances are easily retrievable starting from the graphical scale views.

If the imaged object is complicated, three-dimensional reconstruction may proceed piecemeal with both graphical and analytical methods. With complicated object-space geometry, many plan views and elevations can be constructed, each with its own

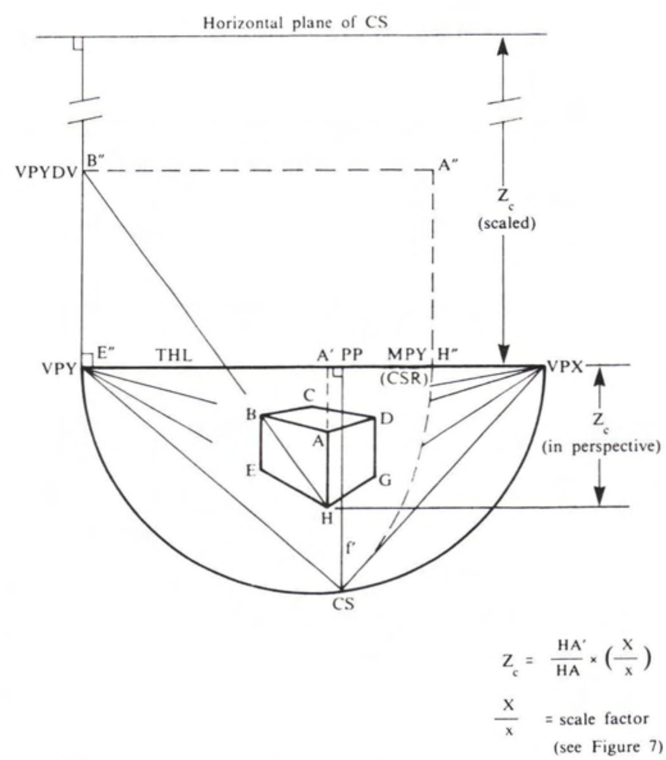


FIG. 8. Two-point perspective true-ratio side view and construction CS coordinate Z_c relative to HA.

scale; the scales are reconciled analytically by referring model dimensions to known object-space dimensions. To acquire the true scale requires knowing the coordinates of at least two object-space points relative to the CS object-space coordinates, or knowing one dimension in a true-ratio view. Many times, the complicated solutions are completed using programmed analytical procedures.

FINDING THE SCALE OF THE SOLUTION ANALYTICALLY

By knowing a dimension in the plan view, it is possible to determine the true scale of the plan view and the CS position. This scaling can provide the object-space X and Y coordinates of the CS. The Z coordinate for the CS is obtained from transferring scale from the plan view to an elevation view and using the ratios as explained

So far, the two-point perspective methods have been strictly graphical with some determination of ratio values. We will now look at a more analytical solution.

For this example, we will use an image with a tilt of 90° (by definition of two-point perspective) and with a swing of 180° (given unformatted imagery). The azimuth angle (a) has been determined from a graphical solution. The rotation matrix then reduced to

$$r_{11} = \cos(a), \quad r_{12} = -\sin(a), \quad r_{13} = 0$$

$$r_{21} = 0, \quad r_{22} = 0, \quad r_{23} = 1$$

$$r_{31} = -\sin(a), \quad r_{32} = -\cos(a), \quad r_{33} = 0$$

and the collinearity equations reduce to

$$x_j - x_{pp} = -f' \frac{[(X_j - X_c)\, r_{11} + (Y_j - Y_c)\, r_{12}]}{[(X_j - X_c)\, r_{31} + (Y_j - Y_c)\, r_{32}]}$$

$$y_j - y_{pp} = -f' \frac{[\qquad (Z_j - Z_c)\, r_{23} \qquad]}{[(X_j - X_c)\, r_{31} + (Y_j - Y_c)\, r_{32}]}$$

(j identifies the object point, and c identifies the camera station)

Using the graphical technique previously described to determine the X_c, Y_c, and Z_c values, it is possible to determine two object-space coordinate values for a point on the object, given the third object-space value. Likewise, given X_c and Y_c along with three object-space coordinates of a point, it is possible to calculate the Z_c value from the y_j equation above. This is the beginning of a bootstrapping analytic method known as "using building planes." By treating any one of the coordinates X_j, Y_j, or Z_j as known from previous solutions, the remaining two coordinates can be determined from the two collinearity equations.

The accuracy of the solution is best determined when there are three or more known dimensions. Some of the known dimensions are used independently or in select groupings to provide scale. The remaining known dimensions are treated as unknowns, solved for, and compared with their known values to determine accuracy. In this case, comparing calculated values to known values is a more practical measure of accuracy than using more formal methods to trace the propagation of error. This is because errors here are of many kinds, some of which are systematic and graphical.

CONCLUSION

This paper clearly shows that when two-point perspective geometry exists in a single image, *a priori* geometric knowledge can enable graphical reconstruction of three-dimensional relationships from that imagery. The methods used to reconstruct three-dimensional relationships might seem contrived, but how many times in the past have single images, with no format and two-point perspective, gone unused because such a simple solution was not considered? In the absence of full format, we used a diagonal of known rotation to determine the camera station and therefore the perspective principal point, perspective focal length, and camera rotation angles. The method of diagonals is just one method available for a graphical photogrammetric solution. There are many other methods, and of course each of them is only as accurate as the known values and the user's ability to work the solution. The choice of two-point perspective methods will always be defined by geometric cues in the imagery, and with man-made objects these cues are nearly always present.

ACKNOWLEDGMENT

Thanks are due to C. P. Kelley, of NDHQ — Ottawa, for the many insights he provided in the course of association with one of the authors (JRW). Also, thanks to Sandy Peterson (SAIC) for the graphics prepared for this paper.

REFERENCES

Beamish, J. (Editor), 1984. Close-Range Photogrammetry & Surveying: State-of-the-Art, *Proceedings of a Workshop*, American Society of Photogrammetry, Falls Church, Virginia.

Busby, W. E., 1981. Dimensional Analysis of a Single Terrestrial Photograph Using the Principles of Perspective, *Fall Symposium*, American Society of Photogrammetry, pp. 431–447.

Duda, R. O., and P. E. Hart, 1973. *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York.

Ethrog, U., 1984. Non-metric Camera Calibration and Photo Orientation using Parallel and Perpendicular Lines, *Photogrammetria*, Vol. 39, No. 1, pp. 13–22.

Gracie, G., *et al.*, 1967. Mensurational Procedures in Terrestrial Photography, *Handbook of Analytical Procedures*, Vol. 2, Naval Air Systems Command, Washington, D.C.

Kelley, C. P., 1978–1983. *Some Aspects of Graphical Techniques Involved in the Rectification of Hand-Held Single Frame Photography*, Seminar Notes, NDHQ-Ottawa.

McCartney, T. O., 1963. *Precision Perspective Drawing*, McGraw-Hill, New York, San Francisco, Toronto, and London.

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

Novak, K., 1986. Determination of Approximate Values for Bundle Adjustment, *Proceedings of the Symposium From Analytical to Digital, Commission III, ISPRS*, Lappia House, Rovaniemi, Finland, August 19–22, pp. 530–542.

Slama, C. C. (Editor), 1980. *Manual of Photogrammetry*, 4th Ed., American Society of Photogrammetry, Falls Church, Virginia.

Walters, Nigel V., and John Bromham, 1970. *Prinicples of Perspective*, Whitney Library of Design, an imprint of Watson-Guptill Publication, New York.

(Received 11 August 1986; revised and accepted 4 November 1986)