Chapter 2

Bases of Three-Dimensional Reconstruction

Paul Allard
Jean-Pierre Bianchi
Rachid Aïssaoui


Many biomechanical three-dimensional (3-D) analyses of human movement start with data capture by an imaging device. Still or high-speed cameras, video cameras, or radiographic systems are the most common of these data capture systems. With a minimum of two different perspectives or views, the spatial coordinates of object points or body markers can be determined using photogrammetric principles. Object points are specified because they can be specifically and uniquely identified in each view and cannot be confused among themselves or with adjoining surfaces or landmarks.

Photogrammetry, defined by the American Society of Photogrammetry, is "the art, science, and technology of obtaining reliable information about physical objects and the environment through the processes of recording, measuring, and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena" (Wolf, 1983). Within photogrammetry, metric and interpretative photogrammetry are two distinct areas. Metric photogrammetry consists of precise measurements from photos and other sources to determine the relative locations of points or markers. Interpretative photogrammetry consists of using systematic analysis to recognize and identify objects. Automated gait-tracking systems involve both the recognition and identification of object markers, as well as their three-dimensional reconstruction.

This chapter covers basic 3-D reconstruction techniques. Classical methods as well as two-step reconstruction techniques are reviewed. The most frequently used of the latter is the direct linear transformation (DLT) method. Sources of errors from different camera types, lenses, and calibration objects are also discussed.
Three-dimensional analysis requires several coordinate systems. In a planar analysis, there is a fixed, or laboratory, coordinate system (FCS), like that shown in Figure 2.1, from which the body marker coordinates are calculated. The axes of the FCS should be labeled unambiguously to avoid confusion with the local coordinate systems, which define joint movement. The origin can be located at any practical point in the laboratory.

A local, or relative, coordinate system (LCS) may be located on a moving body segment. For example, if a kinematic description of the upper limb is sought, an LCS can be located on the shoulder even though the entire upper limb is in motion. Then the body marker coordinates are calculated relative to those of the shoulder. The marker velocities and accelerations are also relative to the shoulder and do not take into account the absolute displacement of the shoulder with respect to a fixed coordinate system located somewhere in the laboratory.

In a three-dimensional reconstruction, two or more different views of the subject are required. These views can be photographed, video or camera frames, and so on. Each view must have its own camera coordinate system (CCS) based in the imaging device.

Figure 2.2 illustrates four camera coordinate systems (CCSs) in which the optical axis defines the Z direction and the X and Y axes are the horizontal and the vertical, respectively, when the camera image is leveled. The origins, Oi (where i is the camera number), are located at the focal point of each camera.

Figure 2.1 Stick diagram of a person walking from right to left. The origins of the FCS can be located at any convenient point in the laboratory. The axes should be identified according to the right-hand rule. The vertical and the direction of progression correspond to the Z and Y axes, and the X axis is in the direction perpendicular to the ZY plane.
Figure 2.2 A four-camera system where the image of all the body markers are identi-
fi ed on the CCS located at the focal point of each camera.

The projection of these coordinate systems onto the image forms the image
coordinate system (ICS) whose origin is the principal point, which is aligned
with the lens focal point. Assuming the image is flat, one axis is eliminated and
the other two axes are labeled u and v, respectively, to avoid confusion between
the image and the camera coordinate systems.

RECONSTRUCTION TECHNIQUES

Once the coordinate systems have been identified, equations are used to determine
the object point coordinates from their image coordinates. (In gait analysis, the
object points are either passive or active markers that identify specific anatomical
landmarks.) These equations relate the external parameters (i.e., the spatial orien-
tation of the camera) and the internal parameters, such as lens characteristics or
camera type, to the object point or marker’s position. Equations and solution
methods have been developed for both fixed and mobile camera setups. Most
three-dimensional gait measurements are carried out with cameras fixed and
facing the subject or patient. In aerial photography, or in machine vision systems
such as self-guided vehicles, the cameras are mounted on a carrier and move along with it. The camera displacement is then taken into account in the calculation.

The standard photogrammetric procedure is based on the exact and precise knowledge of all the internal and external camera parameters. An error in any one of these parameters results immediately in an erroneous coordinate value. Nevertheless, these techniques are extremely good when using metric cameras designed for stereophotogrammetry and when each camera’s spatial position and attitude have been determined accurately.

Using nonmetric or off-the-shelf cameras whose internal and external parameters have been well estimated can yield acceptable results. Because the internal and external parameters have been obtained by an indirect means, this approach is called an *implicit* method as opposed to the direct, *explicit*, method.

**Fixed Cameras**

Three-dimensional reconstruction techniques for fixed cameras can also be used with more than two camera views. The standard technique is based on using two images; any additional images are used to create additional camera pair combinations. Thus, three images form three pairs, whereas four images make six pairs, and so on. This applies to still photography as well as high-speed cinematography or videography. With videography, the image pairs must be synchronized beforehand. However, a single moving camera can also be used if the object is relatively stationary, as in serial photography.

One of the guiding principles of photogrammetric analysis is the collinearity condition. This condition, illustrated in Figure 2.3, requires that an object point

![Figure 2.3](image)

*Figure 2.3* The basic collinearity condition required for 3-D reconstruction is shown by the lines joining markers A, B, and C to their corresponding image crossings at the camera’s focal point.
or marker point and its corresponding image form a straight line that passes through the focal point. (The simplest example of this is the pinhole camera.) If this condition is met, then geometric and trigonometric relationships can be applied to determine the spatial position of the object or marker. In Figure 2.3, the object points, A, B, and C lie on the same straight lines as their corresponding image points, a, b, and c. The collinearity condition equations are developed from similar triangle relationships and are fully detailed in Wolf’s (1983) textbook on photogrammetry.

Equally important is the coplanarity condition, which requires that an object point, A, its corresponding image pair, a and a1, as well as the cameras’ focal points, f1 and f2, lie in a common plane (see Figure 2.4). Whereas the collinearity condition relates the object point to its image, the coplanarity condition links the images (at least two) to the object’s point. With these two conditions fulfilled, the spatial coordinates can be analytically reconstructed from image representations.

**Classical Reconstruction Techniques**

Most of the implicit techniques were derived from classical techniques. A good understanding of the basic approach leads to better comprehension of current practices. Among the many 3-D equations derived from the collinearity condition, the basic formulas of aerial photogrammetry are similar to those used in human motion analysis. The difference is that the optical axis is usually horizontal in a human motion laboratory rather than vertical as in aerial photography.

![Figure 2.4](image)

The coplanarity condition is illustrated in that each body marker lies in a common plane formed by its corresponding image positions and the focal point of each camera.
A stereo-photogrammetric setup is shown in Figure 2.5, with typical distances given for illustrative purposes. The important assumptions are that the external camera parameters are known and that the film planes lie in the same plane. This implies that their optical axes are parallel and perpendicular to the film plane. The \( u_0 \) and \( v_0 \) axes originate at the film principal point, \( P \). Each camera’s optical axis passes through the focal point and is aligned with the respective camera’s principal point. The FCS has been arbitrarily set at the focal point of Camera 1 at \( O_1 \), and the cameras are separated by a base distance, \( B \), of 1.200 m. The focal length, \( C \), is 0.300 m. In this example, the coordinates of an object point, \( A (0.600, 1.300, 0.300) \), expressed in meters with respect to \( O_1 \), are known a priori. They can be determined analytically by applying the following equation (Hallor, 1960).

The equations are

\[
X = (B \cdot u)/p, \quad (2.1)
\]

\[
Y = (B \cdot v)/p, \quad (2.2)
\]

\[
Z = (B \cdot z)/p, \quad (2.3)
\]

where \( p = u_0 - u \) and is used to correct for parallax in the \( O \) coordinate system. Using these equations, the object coordinates are as follows:

\[
X = (1.200 \cdot 0.140)(-0.140 - 0.137) = 0.606 \text{ m}
\]

\[
Y = (1.200 \cdot 0.300)(-0.140 - 0.137) = 1.300 \text{ m}
\]

\[
Z = (1.200 \cdot 0.690)(-0.140 - 0.137) = 0.299 \text{ m}
\]

**Figure 2.5** Basic stereo-photogrammetric setup with the optical axes parallel to each other. The camera base is 1.200 m and the focal length is 0.300 m. Colinearity and coplanarity conditions are met. All dimensions are expressed in meters.
Although the mathematics are quite simple, the technique requires metric cameras accurately positioned. Ayoub, Ayoub, and Ramsey (1976) describe in detail their photogrammetric system and supporting assembly as well as the correction procedure to determine the principal distance, the principal point, and lens distortion.

Today's low-distortion lenses are readily available for stereophotogrammetry with high-speed and video cameras. But we recommend you check the lens for distortion errors. First ask whether the manufacturer can supply you with the lens characteristics. Alternatively, you can do your own calibration with an acceptable accuracy using a sheet of graph paper taped to a granite measuring table. These tables are usually flat to ±2 units. fiducial marks on the graph paper are used to test for lens distortion. The granite table is positioned vertically with the use of a good quality level and set as perpendicular as possible to the camera's optical axis. Although the table can be slightly oblique, this bias should be constant.

Equations 2.1 to 2.3 can be expressed in a more general form:

\[
\begin{bmatrix}
    u - u_0 + d_u \\
    v - v_0 + d_v \\
    -C
\end{bmatrix}
= \lambda[M]
\begin{bmatrix}
    X - X_0 \\
    Y - Y_0 \\
    Z - Z_0
\end{bmatrix}
\]

where \( u \) and \( v \) are the image coordinates, \( u_0 \) and \( v_0 \) are the principal point image coordinates, \( d_u \) and \( d_v \) are the \( u \) and \( v \) errors, \( C \) is the principal distance of the camera, \( \lambda \) is the linear scale factor, \( M \) is a \( 3 \times 3 \) transformation matrix from the image to the laboratory coordinate system, \( X, Y, \) and \( Z \) are the object laboratory coordinates, and \( X_0, Y_0, \) and \( Z_0 \) are the focal point coordinates in the \( X, Y, \) and \( Z \) system. It is from this general form that many implicit three-dimensional equations are derived.

Implicit Reconstruction Techniques

Close-range photogrammetry, or the application of photogrammetry where object point distance is less than 308 m (1,000 ft), became possible with the emergence of suitable instrumentation, methodologies, and techniques for data reduction (Duran, 1975). The departure from metric to nonmetric cameras using new analytical representations for determining the camera's internal and external parameters is an important development in photogrammetry to those involved in the study of human movement. The immediate benefits are in the low cost and readily availability of off-the-shelf cameras, which are not suitable for the classical reconstruction approach.

Today, implicit techniques have proliferated. Many have been developed by mathematicians, expansion, and simplification of Equation 2.4. The classical equation is totally transformed into a series of unknown parameters, which are functions of constants and object point coordinates. The contours are complex mathematical relations involving one or more camera parameters. To determine these parameters experimentally, a calibration device consisting of accurately measured object points is used to determine the unknowns in a two-step approach.
In the first step, the calibration device is photographed and the image and real object point coordinates are used to determine the unknown constant values in the mathematical expression. This is a form of camera calibration. Then, the calibration device is replaced by the unknown object points or body markers to be measured. Their coordinates are determined from the image's object coordinates and the previously calculated analytical constants. The number of calibration points required to solve the unknown parameters vary according to the implicit technique applied.

Several excellent three-dimensional techniques have been published in the field of biomechanics of human movement. To review them is outside the scope of this work. Generally, the error varies between 0.01 mm and 5 mm. However, the reported values are not always based on the same definition of error.

Among these three-dimensional reconstruction methods, the most widely applied and discussed probably is the direct linear transformation (DLT) technique, developed by Marzari (1975). It has been used with many types of imaging devices. Stokes, Bigelow, and Mooreland (1987) applied the DLT in an X-ray photogrammetric technique to measure sciotic spines, while Allard, Duhmtré, Labelle, Murphy, and Nagata (1987) used still cameras to determine the location and orientation of the ankle and subtalar joint axis of rotation. The DLT errors associated with the use of high-speed photography (Shapiro, 1978) and videography (Lemos, Allard, & Murphy, 1990) have been assessed.

The DLT method is based on the classical technique and is expressed as

\[ u + \Delta u = \frac{L_1 X + L_2 Y + L_3 Z + L_4}{L_5 X + L_6 Y + L_7 Z + L_8} \]

\[ v + \Delta v = \frac{L_9 X + L_10 Y + L_11 Z + L_12}{L_13 X + L_14 Y + L_15 Z + L_16} \] (2.5)

where \( u \) and \( v \) are the image coordinates and \( \Delta u \) and \( \Delta v \) are the image coordinate correction for lens distortion. The object point coordinates are \( X, Y, \) and \( Z \), whereas the constants \( L_i \) to \( L_{16} \) are the DLT parameters, which define the camera position and orientation as well as the camera internal parameters and linear lens distortion factors. See Marzari (1975) for a detailed description and two Bioench-I (E-mail) files, which are available to the membership of Biomch.1. with the following command(s) to ListServ@hean.bitnet or to ListServ@nic.sunetnet.nl: SEND DLTDSP README BIOMCH.1. and SEND DLTDSP FORTRAN BIOMCH.1.

In the example based on the classical reconstruction technique (Figure 2.5), the optical axes must be parallel to each other. If they are not parallel, Equation 2.4 cannot be applied, and new sets of equations must be developed to correct for angled cameras. With the DLT equations, the camera can converge, but care must be taken so that the relationship between the convergence and the overlap angles defined in Marzari (1975) is respected to minimize any reconstruction error. This is not always possible in a laboratory environment, and a compromise between camera setting and an acceptable error must be reached.

Equation 2.5 has been implemented on a few commercial video-based kinematic systems to track body markers. For example, they have been applied to the
study of normal and pathological gait. Figure 2.6, a through c, and Figure 2.7, a through c, illustrate two different representations of a normal gait pattern where the corresponding stick diagrams are shown (a) for the plane of progression, (b) the frontal plane, and (c) from above.

Moving Cameras

Moving cameras increase the complexity of calibration and the demands placed on the reconstruction technique. When video-cameras are mounted on a moving vehicle, the computer image analysis and 3-D reconstruction techniques push the computing devices to their extreme limits. Simpler applications can be used to study human movement. For example, the camera can be allowed to rotate about a vertical axis while tracking a subject. The tilt-angle (vertical) and the pan-angle (horizontal) can be obtained from the projection of the reference markers' images in the projection plan of the cameras. De Groot, de Koning, and van Ingen Schenau (1989) modified the technique developed by Dapena, (continued)

Figure 2.6 Overlaid stick diagrams for a person walking from left to right as seen in (a) the sagittal plane, corresponding to the generalized direction of progression, (b) the frontal plane (see next page), and (c) the horizontal plane (see next page). Raw Figures courtesy of Motion Analysis Corporation.
Figure 2.6 (continued)
Figure 2.7 Spaced stick diagram for a person walking from left to right as seen in (a) the sagittal plane, corresponding to the general direction of progression, (b) the frontal plane, and (c) the horizontal plane. 

Note: Figures courtesy of Motion Analysis Corporation.

Haman and Miller (1982) for handling this effect. This technique requires precalibration to determine the transformation matrix.

De Haan and den Brinker (1988) outline the important factors that influence measurement accuracy for subject tracking. Continuous tracking caused some blurring of the reference markers. This was related to both exposure time and marker size. Additionally, the position of the cameras relative to each other can generate substantial error.

Tzai (1987) reported a new 3-D camera calibration technique for machine vision metrology using off-the-shelf TV cameras and lenses. This two-stage
method aims at efficient computation of the camera’s internal and external parameters, which, within limits, can be processed in real time using standard video cameras. Efforts are under way to improve existing 3-D reconstruction techniques or to develop new ones to account for moving cameras. Among the problems that have been identified are the relative camera positioning and computation time limitations. Viable techniques are still needed, but there are ways to cope with the errors that arise.

MINIMIZING RECONSTRUCTED COORDINATE ERROR

Each measurement has an associated error that makes it more or less inaccurate. Bertil Hallert (1960) states that “true” errors are fiction because exact values are seldom known. Thus, it is preferable that the term reference, rather than exact or true, be used in this context. Many factors can influence the quality of the reconstructed coordinates. The image ones are the imaging device, marker identification, and the camera setup and calibration. Before attempting any discussion on these sources of error, a few words of caution about the reporting of the reconstructed coordinate error. Errors are given in the literature as standard deviations, root mean squares, percentages, estimates, and so on. These terms must be carefully defined and used consistently.

Error Definition Terminology

One should distinguish between accuracy and precision. The former refers to systematic differences between the reference (true) and measured values. The latter refers to the repeatability with which a measured value can be obtained. These can be thought of as systematic and random error components, respectively. We recommend that only the root mean square (RMS) error be given and that it be expressed as a fraction (percentage) of the mean camera base-to-object distance. The RMS error is expressed as

$$E_{RMS} = \frac{\sqrt{\sum (x_i - \bar{x})^2}}{N}$$

(2.6)

where $N$ corresponds to the number of observations, $x_i$ is the reference value, and $x$ is the error values. Furthermore, it must be applied only to object point coordinates that have not been used to determine the analytical constants.

To illustrate the difference between the mean error, the absolute mean error, and the RMS error, an example is given for 12 measured and reconstructed markers obtained by one of the authors. Table 2.1 shows that the mean value is very close to $x_0$ as expected. There is a very small bias of about 0.11 mm in the instrumentation and a random error of about 0.25 mm in each coordinate. The mean error is inappropriate for expressing the instrument accuracy, because the negative errors are cancelled by the positive one. Here the absolute mean error
<table>
<thead>
<tr>
<th>Marker</th>
<th>$\Delta X$ (mm)</th>
<th>$\Delta Y$ (mm)</th>
<th>$\Delta Z$ (mm)</th>
<th>Absolute difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>-0.21</td>
<td>0.18</td>
<td>0.43</td>
<td>0.51</td>
</tr>
<tr>
<td>02</td>
<td>-0.18</td>
<td>0.30</td>
<td>-0.20</td>
<td>0.40</td>
</tr>
<tr>
<td>03</td>
<td>0.05</td>
<td>-0.26</td>
<td>-0.02</td>
<td>0.27</td>
</tr>
<tr>
<td>04</td>
<td>-0.09</td>
<td>0.18</td>
<td>0.52</td>
<td>0.28</td>
</tr>
<tr>
<td>05</td>
<td>-0.07</td>
<td>-0.08</td>
<td>-0.34</td>
<td>0.36</td>
</tr>
<tr>
<td>06</td>
<td>0.28</td>
<td>-0.23</td>
<td>-0.45</td>
<td>0.58</td>
</tr>
<tr>
<td>07</td>
<td>0.19</td>
<td>0.11</td>
<td>0.15</td>
<td>0.27</td>
</tr>
<tr>
<td>08</td>
<td>0.14</td>
<td>-0.32</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>09</td>
<td>0.07</td>
<td>0.03</td>
<td>-0.67</td>
<td>0.28</td>
</tr>
<tr>
<td>10</td>
<td>-0.04</td>
<td>0.27</td>
<td>0.27</td>
<td>0.38</td>
</tr>
<tr>
<td>11</td>
<td>-0.35</td>
<td>0.33</td>
<td>0.09</td>
<td>0.49</td>
</tr>
<tr>
<td>12</td>
<td>-0.00</td>
<td>-0.02</td>
<td>-0.14</td>
<td>0.14</td>
</tr>
</tbody>
</table>

**Mean error**

- $\Delta X$: 0.02 mm
- $\Delta Y$: 0.04 mm
- $\Delta Z$: 0.10 mm

**Standard deviation**

- $\Delta X$: 0.05 mm
- $\Delta Y$: 0.23 mm
- $\Delta Z$: 0.35 mm

**Absolute error**

- $\Delta X$: 0.14 mm
- $\Delta Y$: 0.19 mm
- $\Delta Z$: 0.29 mm

**RMS error**

- $\Delta X$: 0.17 mm
- $\Delta Y$: 0.26 mm
- $\Delta Z$: 0.31 mm

*Note: Values were calculated from the measured and reconstructed data using the DLT method.*

(0.82 mm) and the RMS error (0.44 mm) are three and four times greater, respectively, than the mean error. The RMS error is a conservative estimate of the instrument accuracy.

Often in photogrammetry, the error is expressed as a ratio of the camera base-to-object distance. In this case, the RMS camera base-to-object ratio would be $1/2.931$. However, if a different lens (12-mm or a zoom lens) had been used instead of the 8-mm lens, the error would have been different even with a similar camera position. Increasing the focal length of the lens is comparable to bringing the camera closer to the object. Using the same camera and lens, the error varies nearly linearly with the camera base-to-object distance.

Because the camera base-to-object distance can vary from one camera setup to another and considering that the focal distance used can be easily modified, it is better to express the error as a proportion of the field of view. This has the effect of lowering the ratio, because the width of the field of view is usually less than the camera base-to-object distance. For the above case, the ratio is now $1/4.141$, shown in Table 2.2. The ratio is still $k$ constant, because the field of view is in proportion to the camera base-to-object distance, shown in Figure 2.8. Triangle ABC is similar to A'BC', thus $AB/A' B'$ is the same as $AC/A'C$. How should the field-of-view length be defined? Is it the vertical, horizontal, or
<table>
<thead>
<tr>
<th>Distance</th>
<th>Mean</th>
<th>Absolute error</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera-base-to-object of 1,298.3 mm</td>
<td>1:1,803</td>
<td>1:3.509</td>
<td>1:2.951</td>
</tr>
<tr>
<td>Field of view</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal (410 mm)</td>
<td>1:3.727</td>
<td>1:1.108</td>
<td>1:0.92</td>
</tr>
<tr>
<td>Diagonal (410 by 205 mm, 458.4 mm)</td>
<td>1:4.167</td>
<td>1:1.239</td>
<td>1:1.042</td>
</tr>
<tr>
<td>Spatial diagonal (410 by 205 by 205 mm, 502.2 mm diagonally)</td>
<td>1:4.765</td>
<td>1:1.357</td>
<td>1:1.141</td>
</tr>
</tbody>
</table>

**Figure 2.8** Similar trigonometric relationship between different camera positions.
diagonal distance? Although any one of these is correct, the diagonal yields a better appreciation of the field of view, because it includes the height and width. However, because they are all related to the camera base-to-object distance, which can easily be measured compared to a camera’s field of view, the mean camera base-to-object distance should be used. The diagonal of the calibration volume represents only a portion of the camera’s field of view, and its use would yield a low ratio.

Any reference marker used previously in the calculation of the analytical constants will generally yield the smallest error, because the analytical parameters have already been optimized for such points. (That the “error” calculated using such reference markers is minimal only reveals the reliability of the optimization process.) Consequently, only the remaining reference marker coordinates should be used in the error calculation. Marzban’s (1975) DLT computer program prints out the standard error or DLT error on the 11 DLT parameters. When we attempted to correlate the DLT error to the absolute error (i.e., the means of the absolute differences between the reference and the measured coordinates), we found that the DLT error was always significantly lower (André, Danseuse, & Allard, 1980). In a simulation trial, when the reference marker error was increased from 0 mm to 5 mm, the DLT error increased from 0 mm to 0.4 mm, whereas, the absolute error reached a maximum of 1.5 mm with peak error close to 3.5 mm.

In summary, the RMS error should be calculated from marker coordinates that have not been used in the calibration process, and the ratio of the RMS or the absolute error should be expressed in relation to the mean camera base-to-object distance.

**Camera Types and Lens Distortion**

The principal components of a stereo-photogrammetric optical system are the camera used, either metric or nonmetric, the film unflatness, and the lens. Each of these contributes to the measurement error, but basic three-dimensional reconstruction models do not correct for these sources of error.

Implicit techniques, such as the DLT, replace the painstaking steps of camera calibration. Karara and Addey-Abir (1974) compared the accuracy of the DLT method using a metric camera and four types of nonmetric cameras. The nonmetric cameras were a Honeywell Pentax Spotmatic, Crown Graphic, Hasselblad 500 C, and a Kodak Instamatic 154 camera; the metric camera was a Hasselblad 70. The standard deviation RMS error was calculated for reference points located at camera base-to-object distances of 4.0 m and 5.5 m. The RMS values were always smaller for the closest camera positions. The largest standard deviation of the RMS error was 3.1 mm (obtained with the Kodak, Instamatic camera).

representing a ratio of about 1:1.780 with respect to the camera base-to-object distance. The RMS error with the other nonmetric cameras was 55% to 75% smaller than the Instamatic camera error. These were too very much different from the metric camera error. The ratios calculated with respect to the camera base-to-object distance varied from 1:1.774 (Instamatic) to 1:5.200 (Pentax).

Our own work, using Nikon FE camera filled with Nikkor 55-mm macro lenses, yielded an average absolute error of 0.33 mm (1:1540). Our accuracy
was limited by the digital graphic tablet error of 0.4 mm and the manual digitiza-
ization process.

Frazer (1982) was able to reach a precision of 1:10,000 or better with nonmetric
 cameras and a DLT-type method. A series of multistation photogrammetric
 adjustments was made using the concept of self-calibration with additional parame-
ters set for each photograph. Among these are the photo-invariant factors,
which take into consideration film deflection peculiar to each photograph but
assume a degree of film flatness that does not vary from photo to photo. According
 to Frazer (1982), lack of film flatness appears to be the most significant limiting
factor on the attainable accuracy of nonmetric camera reconstruction techniques,
especially for large-format cameras and short focal length lenses.

Reconstruction techniques involve the calculation of several constants related
to the camera's internal and external parameters, lens correction factors, and other
characteristics. Some of these constants are an integral part of the reconstruction
 technique, whereas others are additional correction factors added to the photo-
 grammetric model. For example, a basic photogrammetric model assumes no
 lens distortions. If these are included, then specific analytical representations of
the type of distortion must be added. Yet if this is done when little or no lens
distortion exists, then random error or bias is introduced.

The Aff DL T three-dimensional algorithm contains 22 parameters. The first
1 are associated with the internal and external parameters and with linear lens
 distortion factors. A polynomial of up to the seventh order can be added to take
into consideration the symmetrical lens distortion, increasing the total number
of coefficients to 16. The image refinement components \( \Delta s \) and \( \Delta S \) from Equation
2.5 can be expressed as

\[
\Delta S = \Delta u + 2 \Delta a \times + \Delta r \times + \Delta (k_r x^2 + k_y y^2) + \Delta (k_r x^2 + k_y y^2)
\]

\[
\Delta s = \Delta a + \Delta \times + \Delta r \times + \Delta (k_r x^2 + k_y y^2) + \Delta (k_r x^2 + k_y y^2) + \Delta (k_r x^2 + k_y y^2)
\]

(2.7)

where the \( a \) are constants that reflect the linear components of lens distortion
and film deformation. These are incorporated into the \( L_1 \) to \( L_{16} \) coefficients,
 whereas \( r \) is the length of the radial vector from the point of symmetry to the
point under consideration \( (u', v') \). The first five additional unknowns, the \( k_r \), are
the symmetrical lens distortion coefficients.

The asymmetrical lens distortion reflects the distortion caused by the decen-
tering of lens elements and accounts for the selection of a point other than that
of symmetry as referred. Conrady's model (1919) is used to express asymmetry-
lic lens distortion. Equation 2.7 becomes

\[
\Delta s = \Delta a + \Delta \times + \Delta r \times + \Delta (k_r x^2 + k_y y^2 + k_x x^2 + k_y y^2) + \Delta (P_x x^2 + 2P_x x + \Delta P_x x^2 + \Delta P_x x^2)
\]

\[
\Delta s = \Delta a + \Delta \times + \Delta r \times + \Delta (k_r x^2 + k_y y^2 + k_x x^2 + k_y y^2) + \Delta (P_x x^2 + 2P_x x + \Delta P_x x^2 + \Delta P_x x^2)
\]

(2.8)

where \( P_x \) and \( P_y \) are the asymmetrical lens distortion coefficients.
Four of the 22 coefficients can be added to Equation 2.8 to account for the nonlinear component of film deformation leading to the complete image refinement model with 22 coefficients.

\[ \Delta u = a_1 + a_2 \mu + a_3 \nu + a_4 \mu \nu + a_5 \rho \mu + a_6 \rho \nu + a_7 \rho \mu \nu + P_1 \rho + 2P_2 \rho \nu + P_3 \rho \mu \nu \]

\[ \Delta v = a_8 + a_9 \mu + a_{10} \nu + a_{11} \mu \nu + a_{12} \rho \mu + a_{13} \rho \nu + a_{14} \rho \mu \nu + P_4 \rho + 2P_5 \rho \nu + P_6 \rho \mu \nu \]  

(2.9)

Thus, the \( \Delta u \) and \( \Delta v \) of Equation 2.5 must be substituted for by their corresponding values in Equation 2.9.

Using video cameras, the film is replaced by a detector and electronic system, which should produce a linear relationship between a light spot’s location on the image plane and the resulting data. A detailed mapping of video image-plane errors was carried out by Antonsson and Mann (1989), who used 12,000 points in an effort to calibrate the positional accuracy over the camera’s entire field of view at a 3-m range. This extremely large number of reference points was generated by a 2-m-long channel section onto which 30 reference markers were positioned along the bar and measured to an accuracy of less than 1 mm. The rod was mounted on a precision motor-driven and computer-controlled rotary table with its axis of rotation aligned with the optical axis. The table was rotated in 0.9° increments to obtain 400 steps around the circle. Plus of nonradial and sociocentrefugal error contours for a video camera are far from being purely radial, as shown in Figure 2.9, a and b.

**Object Calibration and Reconstruction Error**

There are many sources of error, but probably the most important and most often neglected is related to the quality and the care given to the measurement of the calibration object. It is impossible to calibrate a device or an instrument to an accuracy greater than that of the standard to which it is compared. Thus Doebelin (1975) recommends use of a standard calibration 10 times as accurate as the accuracy required of the device or instrument being calibrated.

In photogrammetry, this has further implications because we use an object that has been physically measured once to calibrate the cameras’ internal and external parameters. Consequently, the accuracy of the instrument used in calibrating the referenced object is crucial because it also affects the accuracy of the 3-D reconstruction. Figure 2.10 conceptually summarises the sources of errors that can be carried in the different steps toward obtaining the reconstructed coordinates. Often, the coordinates of the reference markers on the calibration object are determined using a standard measuring tape that has an accuracy of 0.5 mm; the reader must judge the reported accuracy of the reconstructed 3-D coordinates. Some manufacturers of video-based systems supply laboratory-calibrated devices, which should contribute to error reduction.

Besides the attention given to the measurement of the calibration object, there are a number of other factors that influence the overall accuracy. These include the number and choice of reference markers used in the reconstruction technique.
Figure 29: Plot of (a) isoradial and (b) isocurvature correction contours.

the shape of the calibration object, the relative position of the camera, and the reconstruction algorithm.

The minimum number of points selected for solving the unknown parameters in the reconstruction algorithm varies according to the technique used. For the DLT, the image coordinates of at least 6 reference markers are required. Shapiro (1978) recommends the use of 12 to 20 reference markers. If more than 6 are used, a least-squares fit of the parameters is performed. The error is not dramatically influenced by the number of points, as long as the calibration object is measured with good accuracy. Leroux, Allard, and Murphy (1990) measured marker position to 10.02 mm within a calibration frame of about 0.027 m² and with a camera base-to-object distance of 0.600 m. With 8 and then 21 reference markers, the magnitude of the absolute error was 0.41 mm and 0.34 mm, respectively.

Caution must be exercised when increasing the number of points to reduce the reconstruction error. A least-squares approach should yield a best fit, and the use of slightly more reference markers than the strict minimum is encouraged. However, although least-squares fit is used to minimize the random error, it should not be used to correct for an experimental bias that results from a poorly calibrated reference object. The error must be estimated not only on relative lengths, such as the distance between two markers, but also on the absolute coordinates of the markers themselves or the length of the vector originating at the laboratory coordinate system and the marker.

The distribution of the reference markers within the calibration is also important. Leroux et al. (1990) compared the errors after calibrating with additional
markers biased toward the front, bottom, or right side of a reference cube. Initially, there were eight markers, one on each corner. Seven additional markers were added. Table 2.3 summarizes the results, which show that a biased distribution of reference markers roughly doubles the error, especially in the direction perpendicular to the film plane (Z direction).

Once the photogrammetric system is calibrated, then object marker positions can be determined. To avoid extrapolation errors, these should fall within the calibration frame. Table 2.4 presents the absolute error in the coordinates of object markers lying to the back, to the left, or higher than the calibration.

### Table 2.3 Absolute Error (Standard Deviation) for a Nonuniform Marker Distribution

<table>
<thead>
<tr>
<th>Marker position</th>
<th>Ex (mm)</th>
<th>Ey (mm)</th>
<th>Ez (mm)</th>
<th>Magnitude $\sqrt{Ex^2 + Ey^2 + Ez^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottom</td>
<td>0.23</td>
<td>0.18</td>
<td>0.46</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.18)</td>
<td>(0.36)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>Right</td>
<td>0.22</td>
<td>0.23</td>
<td>0.48</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.22)</td>
<td>(0.36)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>Front</td>
<td>0.23</td>
<td>0.23</td>
<td>0.47</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.22)</td>
<td>(0.40)</td>
<td>(0.49)</td>
</tr>
</tbody>
</table>

### Table 2.4 Absolute Error (Standard Deviation) in Coordinates of Object Markers

<table>
<thead>
<tr>
<th>Marker position</th>
<th>Type</th>
<th>Ex (mm)</th>
<th>Ey (mm)</th>
<th>Ez (mm)</th>
<th>Magnitude $\sqrt{Ex^2 + Ey^2 + Ez^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Back</td>
<td>Intra</td>
<td>0.16</td>
<td>0.17</td>
<td>0.31</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13)</td>
<td>(0.15)</td>
<td>(0.20)</td>
<td>(0.28)</td>
</tr>
<tr>
<td></td>
<td>Extra</td>
<td>0.21</td>
<td>0.33</td>
<td>0.46</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.24)</td>
<td>(0.26)</td>
<td>(0.31)</td>
<td>(0.47)</td>
</tr>
<tr>
<td></td>
<td>Left</td>
<td>0.17</td>
<td>0.21</td>
<td>0.48</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.10)</td>
<td>(0.17)</td>
<td>(0.24)</td>
<td>(0.31)</td>
</tr>
<tr>
<td></td>
<td>Extra</td>
<td>0.47</td>
<td>0.36</td>
<td>1.37</td>
<td>1.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.23)</td>
<td>(0.29)</td>
<td>(0.96)</td>
<td>(1.02)</td>
</tr>
<tr>
<td></td>
<td>Up</td>
<td>0.14</td>
<td>0.26</td>
<td>0.44</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.08)</td>
<td>(0.17)</td>
<td>(0.19)</td>
<td>(0.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.27)</td>
<td>(0.26)</td>
<td>(0.20)</td>
<td>(0.47)</td>
</tr>
</tbody>
</table>

Note: Intra—markers lying within the calibration space. Extra—markers lying outside the calibration space.
frame: The error has doubled or more. Wood and Marshall (1986) found that extrapolation errors were 50% to 100% greater.

Throughout this chapter, two assumptions have been maintained: a) that markers in multiple images are easily identifiable, and b) that they are representative of well-defined body landmarks or articulations. However, in practice, data files must be edited to eliminate noise from interfering light sources, to redefine markers after occlusion or shadowing problems, to interpolate missing data, to correct for marker displacement resulting from skin movement, and so on. Wolting and Huiskes (1990) foresee the time when pattern recognition and artificial intelligence techniques may be more widely applied to data-capture systems for the analysis of human movement.

ACKNOWLEDGMENTS

The financial support by the Natural Sciences and Engineering Research Council of Canada is greatly appreciated, in particular for the equipment and operating grants to carry out the 3-D analysis of human motion. The technical support of Althee Mongrain is also acknowledged.

REFERENCES


