Stereophotogrammetry

7

MEASUREMENT OF BODY SEGMENT MOTION

7.1 STEREOPHOTOGRAMMETRY

H. J. Woltring and R. Huiskes

Since the early 1970's the use of analytical, close-range photogrammetry in biomechanics has been increasingly apparent from the biomechanical literature, following the pioneering work at the turn of the century by Braune and Fischer in Leipzig and the subsequent studies of Bernstein in Russia and of the Berkeley group in the USA. For reviews see Selvik (1974, 1983), Woltring (1984a) and Huiskes et al. (1985).

While photogrammetry started about a century ago as a photography-oriented science, the current definition as listed in the *Manual of Photogrammetry* (Slama, 1980) is:

> Photogrammetry is the art, science, and technology of obtaining reliable information about physical objects and the environment through processes of reading, recording, and interpreting photographic images and patterns of electromagnetic radiant energy and other phenomena.

Thus, nonphotographic approaches which are particularly relevant in biomechanics are expressedly included. Other, more conventional applications include surveying, topography, and remote sensing from space. The term "close-range photogrammetry" is usually reserved for observation distances up to a few 100 m.

In all photogrammetric data acquisition systems, whether photographic, cinematographic, optoelectronic, or Röntgenographic, some type of camera system is used as a direction sensor for incoming radiation. Usually, the sensor is 2-D, but dedicated line scanners may provide 1-D information per sensor. Combination of multiple sensors and/or the assumption of given

subspaces of interest (e.g., a "plane of motion") then yield sufficient information for uniquely determining the physical data being investigated. In biomechanics, these are usually shape and position data.

The smooth nature of the anatomy of the body and the impossibility to derive axial rotation from axio-symmetric contour data necessitate the use of easily recognizable landmarks on the body. In shape determination, a raster of lines (e.g., Huiskes et al., 1985) can be projected onto the body, while in motion studies a number of special markers is often affixed to the body segments. Depending on the nature of the measurement system, these landmarks may be passive (e.g., bright reflectors) or active (e.g., wired IR LEDs), and they may exhibit unique characteristics so as to allow their automatic identification per camera and in time. Of course, wired landmarks are cumbersome, but their identification is considerably facilitated. If the landmarks cannot be identified in hardware (shape, color, time and frequency multiplexing), software pattern recognition procedures and/or human interaction are necessary. This can be a time-consuming effort, and various procedures such as extrapolation/prediction in time and space are used in this context. Once the markers have been identified, camera calibration and 3-D landmark position reconstruction from given image data are standard problems of photogrammetry with strong reliance on (iterative) least-squares calculus.

Efficient camera calibration tends to be rather complicated. Typically, unknown camera parameters such as positions and attitudes with respect to an operationally defined coordinate system are surveyed or reconstructed from observations on a suitable distribution of control points with (more or less) known spatial positions. Here, a tradeoff must be made between the complexity and accuracy of the required calibration rig and of the calibration software package. Some procedures such as the Direct Linear Transformation (DLT) require a complex, 3-D calibration rig with known control point positions throughout the volume of interest; others may require less precise control information at the expense of more complicated software systems.

Once a camera configuration has been calibrated, body segment kinematics can be reconstructed from partial or complete landmark measurements. For complete measurements, the raw image data can be used to reconstruct the individual landmarks for subsequent use in a rigid-body reconstruction procedure. For incomplete measurements, e.g., caused by shadowing effects, photogrammetric data can be combined with known rigid-body constraints. See also Chapter 10 in this book for further aspects of kinematic data processing.

DATA ACQUISITION[1]

Photography and Cinematography

Study of the kinematics of human movement has a long history: see Grieve, et al. (1975) and Woltring (1984a) for methodological reviews. Until the 1970's, cinematography and multiple-exposure stroboscopic photography were the prevailing tools for data acquisition. Typically, small landmarks are affixed to the subject's body, and their spatial positions are recovered by digitizing the image coordinates on a suitable image-data convertor, similar to photogrammetric procedures in cartography and surveying.

Manual conversion of cinematographic data to quantitative form on a frame-to-frame basis is a very laborious task, particularly in the 3-D case with multiple cameras. This situation is alleviated by current digitization facilities in a semiautomatic, computerized way, an approach which is notably followed in impact biomechanics. Typically, an operator identifies the relevant landmarks by means of a lightpen, and a pattern recognition procedure tracks corresponding image points in subsequent image frames. Whenever landmarks are obscured or confused with neighboring landmarks, the operator must intervene.

Video Systems

Since the 1960's, various alternatives for cinematography (including stroboscopic photography) have been developed. Among these, a number of optoelectronic systems have become available both commercially and academically. Here, the digitization of image coordinates is automatic, and some systems also provide hardware identification of multiple landmarks. This distinction may seem trivial, but it appears that determining which image point corresponds to what landmark can be quite cumbersome, especially in complex 3-D movement. The earliest methods were based on video systems (e.g. Winter et al., 1972), where the image coordinates of retroreflective markers or small light bulbs are read by means of a custom-designed interface, and transferred to a recorder or computer for further processing. Identification of corresponding image points in different video frames between cameras and time instants is done via pattern recognition software as in the cinematographic case (e.g. Taylor et al. 1982); in recent systems, some of the pattern recognition is done in hardware (Ferrigno and Pedotti, 1985). Macellari (1983) and Mesqui et al. (1985) have described systems based on single-axis photodiode arrays, in combination with time

---

[1] This section was largely taken from Woltring (1984b).

multiplexed LEDs for automatic identification of multiple markers. The system of Mesqui et al. contains a minicomputer for real-time, 3-D landmark position reconstruction. In Section 7.2, Furnée describes a solid-state video system with 1:32.000 image resolution and 100 Hz sampling frequency.

A commercial and computerized video-system VICON[2] has been available since 1982. It can accomodate up to seven video-cameras, and it uses retro-reflective markers which are affixed to the body, in combination with infrared, stroboscopic illumination. A ring of IR LEDs is mounted around the lens of each camera, and the LEDs give a short light flash at the end of each video scan. In this fashion, image blur caused by fast movement is avoided in the same way as in high-speed photography. The camera lenses contain an IR lowpass filter in order to minimize the influence of background light. However, sunlight and strong incandescent light must be avoided. A comprehensive software package (for use on PDP11 and VAX computers) is provided, allowing 3-D camera calibration, data collection, reduction, and sorting (landmark identification) of raw image data, 3-D landmark reconstruction, smoothing and differentiation, and graphics. Additional data channels are available for synchronized collection of EMG and force-plate data. The number of landmarks can be very high and is mainly limited by the disk transfer rates for data storage. A considerable asset of the system is the passive nature of the landmarks (no wires), and the availability of a comprehensive, well-designed software package. Disadvantages are the processing time required for reduction of the data and for operator-supervised sorting, and the relatively low spatio-temporal resolution (usually standard video: 1:600-1000; 50-60 Hz).

Lateral Effect Photodetector Systems

Another system for multipoint monitoring is based on a completely different principle. The SELSPOT-II system[3] relies on the lateral photoeffect (Lindholm and Öberg, 1974; Woltring, 1975; Woltring and Marsolais, 1980) for determination of the position of a light image on a semiconductor photodetector (see Fig. 1). The incident light causes a photocurrent which divides itself between the lateral contacts on both sides of the detector. If the detector is fully reverse-biased, the relation between the position of a light spot and the signal currents is linear (Woltring, 1975). Multipoint monitoring occurs via time-division multiplexed operation of IR LEDs which are affixed to the subject's body as landmarks, and background light influence is compensated for by operating on the changes in the output

---

[2] VIdeo CONvertor; trademark of Oxford Metrics, Oxford, United Kingdom.

[3] SELective light SPOT recognition, trademark of SELSPOT AB, Partille, Sweden.

currents caused by the flashing LEDs. Spatial resolution is 12 bits (1:4096) per image axis, temporal resolution is 100 μs per LED, and, for enhanced resolution, per camera (a feedback option is provided which controls each LED's light intensity so as to receive sufficient light in each camera); the maximum number of LEDs is 128.

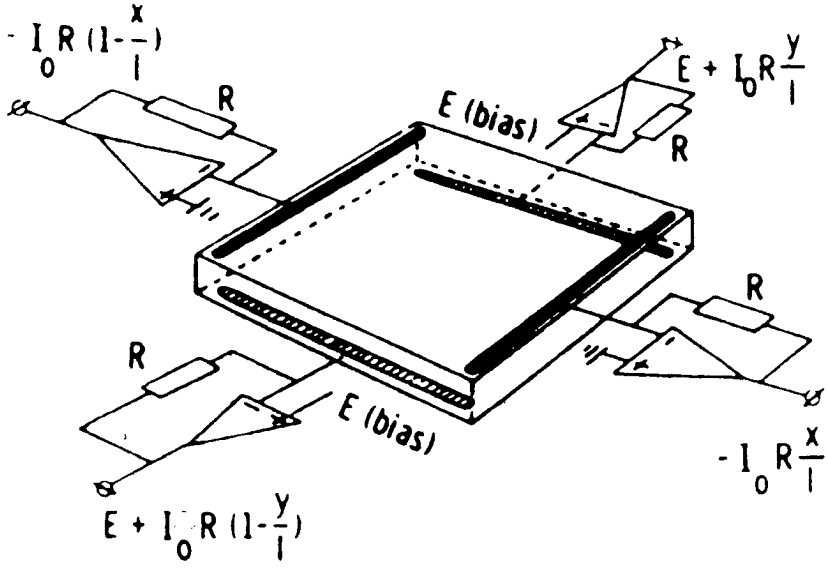

Figure 1. Dual-axis, duo-lateral photodiode for 2-D, linear position detection of the centroid of an incident light distribution (from Woltring, 1975).

A similar system called WATSMART[4] using single-axis detectors with cylindrical optics is also available. Advantages of SELSPOT-II and WATSMART with respect to video systems are the (limited) possibility to use the system in daylight situations, the much higher spatial resolution, the high resolution in time (unless the number of landmarks is very high), and automatic landmark identification. Disadvantages are the use of active (wired) markers and the sensitivity to marker light reflections: since the lateral detector senses the centroid of the incident light distribution, any indirect images due to reflections on neighboring surfaces (adjacent limbs, ground, walls, ceiling) will influence the position estimate for the directly observed light source. The only way to prevent this is to minimize reflectivity by using proper garments, paints, tapestry, and antireflective sprays. The SELSPOT system can be provided with a software package, but this would currently appear not to be as comprehensive as in the case of the VICON system. A number of similar systems of Japanese and American origin are also available, but their utilization in biomechanics seems limited to a few cases only.

---

[4] WATSMART - Trademark of Northern Digital Inc., Waterloo, Ontario, Canada

Optomechanical Scanners

One.other approach is based on scanning mirrors. Güth et al. (1973) and Heinrichs (1974) described a method in which a V-shaped light image is projected onto a rotating mirror. The mirror sweeps the reflected image through space, and the legs of the V repeatedly hit small photodetectors affixed as landmarks to the subject's body. The times at which a photocurrent pulse occurs are converted into 2-D direction information, with a resolution ≥ 1 mm at 8 m distance and 40 Hz sampling frequency. Similar systems are known in robotics and in automatic shape detection systems. A commercial system for biomechanical purposes has recently become available. This CODA-3 system[5] uses colored, retroreflective landmarks on the subject. Two mirrors rotate about vertical axes which are spaced 1 m apart. A third mirror is positioned halfway between these mirrors, and it rotates about a horizontal axis which intersects the two vertical axes. A line of light is projected onto each mirror, and the reflected lines sweep through space because of the mirror rotations. Each mirror scans over an angle of 40 degrees at a rate of 300 Hz. Whenever a retroreflective landmark is hit by a light beam, colored light is reflected back into the scanning mirror. Time and color are determined via splitting optics, diffractive gratings, and optoelectronic circuitry, and these provide unique direction and identification information. Tangential resolution is about 1:16000 over the 40 degrees scanning angle, and this corresponds to about 50 μm per m distance.

CODA-3 incorporates a set of microcomputers for real-time, 3-D landmark position reconstruction. Attainable precision is determined by the fixed geometry: the fixed base of 1 m between the vertical mirror axes entails a quadratic dependence on the distance from the mirrors for the longitudinal coordinate, and a linear dependence for the two other coordinates. Advantages with respect to the SELSPOT-II/WATSMART and VICON systems are the passive nature of the landmarks, the high spatio-temporal resolution, and the real-time availability of 3-D landmark coordinates. Disadvantages are the limited number of landmarks (8-12), due to difficulties in distinguishing the landmark colors, and the fixed stereobase which limits the depth range.

Clearly, the various optoelectronic systems are complementary in terms of the number and of the active/passive nature of the landmarks, spatio-temporal resolution, real-time properties, background light influence, and subject encumbrance. None is superior in all respects, and the choice between them depends on the individual application. A useful criterion for

---

[5] CODA-3, for Cartesian Optoelectronic Digital Anthropometer, trademark of Charnwood Dynamics Ltd, Loughborough, Leic., United Kingdom.

comparing spatio-temporal resolution between systems is the value of $\sigma^2\tau$ where σ^2 is the variance of a reconstructed coordinate in a standardized environment, and τ the sampling interval: see Woltring (1984a, 1984b) and Chapter 10 in this book.

Röntgenographic Systems

Like the optical system discussed above, Röntgenographic systems use electromagnetic radiation as the measurement vehicle; however, lenses are absent, the perspective center is formed by the Röntgen source, and the object of interest is placed between the perspective center and the recording device, usually a photographic plate: the object is, as it were, located inside the camera(s); see Fig. 2. The Röntgen absorption pattern is recorded, and the relevant features are digitized for subsequent processing.

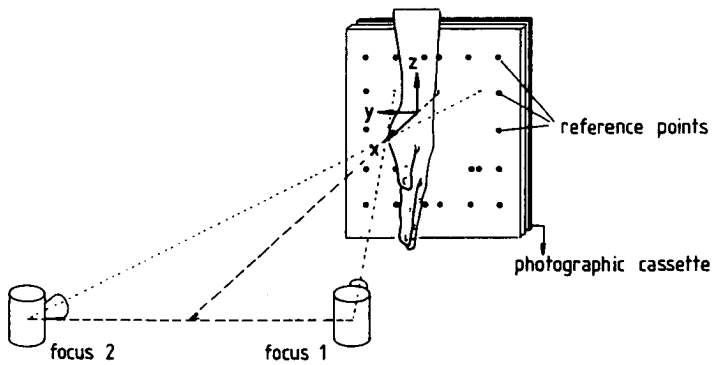

Figure 2. Röntgenstereophotogrammetric measurement configuration (from Woltring et al., 1985).

Also here, the use of artificial markers can substantially enhance the utility of the photogrammetric process, both in vivo (Selvik, 1974, 1983) and in vitro (Huiskes et al., 1985). The use of small tantalum markers (0.5 to 1.0 mm diameter) is relatively easy and when properly implanted into the relevant body parts not detrimental to the live subject. Not only are image contrast and resolution enhanced, but the required Röntgen doses for adequate image exposure can be substantially smaller. While high measurement frequencies in real-time are difficult to realize, movement steps followed by film exposure during in vitro research can be made sufficiently small in order to emulate continuous in vivo movement.

Spatial resolution can be very high, e.g. on the order of 10 μm per image coordinate, yielding reconstruction errors in depth on the order of 25 μm. Major error sources are the finite apertures of the Röntgen foci which produce blurred marker images, and the difficulty for a human operator to accurately determine the centroid of such a blurred image; automatic centroid

detection systems can reduce this error type. Like in all systems with passive, nondistinct markers, identification of these markers can be difficult, and pattern recognition procedures can enhance the routine use of such systems considerably.

Ultrasonic Systems

Although strictly not belonging to the class of photogrammetric systems, a short discussion of ultrasonic equipment seems not out-of-place, in view of the similarity with photogrammetry. While photogrammetric detectors are (usually) direction sensors, ultrasonic digitizers as used in the biomechanics field are distance sensors. The markers consist of small sound generators, e.g., spark plugs, and the traveling time of the sound wavelet to a set of microphones is used to determine the distances of the sound source to the microphones. With similar, nonlinear models as are used in photogrammetry, the two- and three-dimensional positions of the sound sources can be reconstructed (e.g., Brumbaugh et al., 1982). Surveyors rely extensively on distance measurement with laser equipment, often in combination with photogrammetric methods.

PRINCIPLES OF CLOSE-RANGE PHOTOGRAMMETRY

Most photogrammetric systems obey the laws of central projectivity: ideally, all points in the observed volume are projected along straight lines onto a given image plane, and these lines intersect in a common point known as the perspective center, see Fig. 3.

Ideal Cameras

Given an observed point P and its image p, the coordinates $\mathbf{X}_{po}$ of P in the object-space coordinate system E_o and the coordinates $\mathbf{x}_p$ of p in the image coordinate system E_c are related as

$$\mathbf{X}_{po} - \mathbf{X}_{co} = \lambda \mathbf{R}'_{co}\left(\mathbf{x}_p - \mathbf{x}_c\right) \tag{1}$$

where λ is a proportionality constant, $\mathbf{X}_{co}$ the position of the perspective center in E_o (the "camera station"), $\mathbf{R}_{co} = [\mathbf{r}_{co_x}, \mathbf{r}_{co_y}, \mathbf{r}_{co_z}]'$ the attitude matrix of E_c with respect to E_o, $\mathbf{x}_p = (x_p, y_p, 0)'$ the observed image coordinates in E_c, and $\mathbf{x}_c = (x_C, y_C, c)'$ the position of the perspective center in E_c. The projection of the perspective center onto the image plane is called the principal point, with coordinates $(x_C, y_C, 0)'$, and the distance c between these points is called the principal distance.

If $\mathbf{X}_p$ is in a given plane with known coordinates in E_o, the scaling factor λ in Eq. 1 can be eliminated by means of the constraint equation

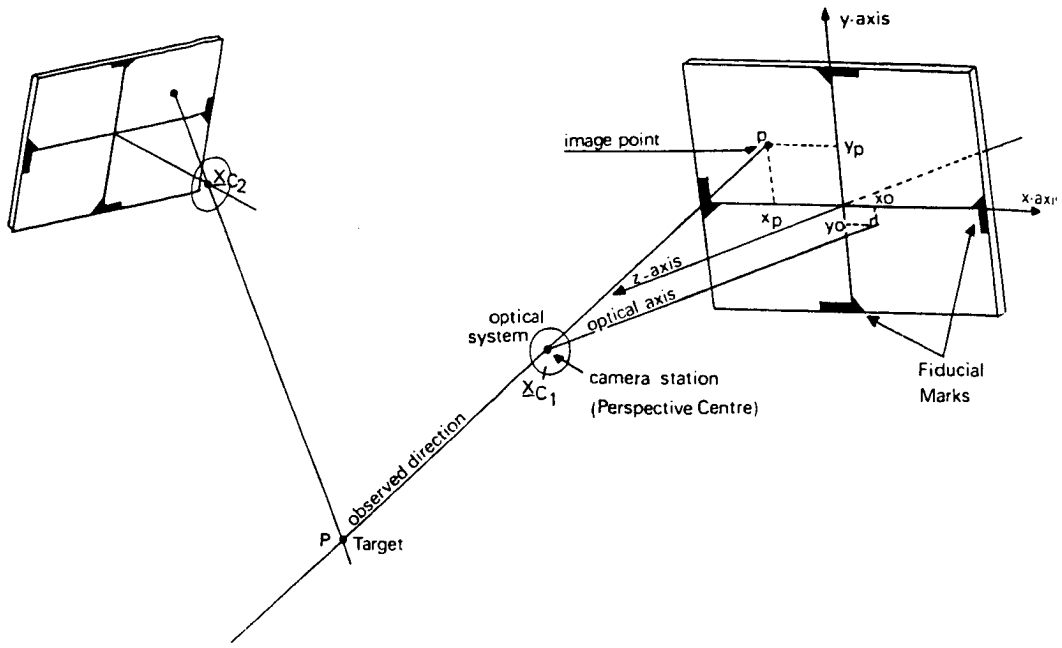

Figure 3. Relation between internal and external camera parameters in target position reconstruction.

$$\mathbf{N}_o' \left(\mathbf{X}_p - \mathbf{X}_o \right) = 0 \tag{2}$$

where $\mathbf{X}_o$ is the position in E_o of some given point in the plane, and $\mathbf{N}_o$ the normal vector onto the plane, with solution

$$\lambda = \mathbf{N}_o' \left(\mathbf{X}_o - \mathbf{X}_{co} \right) / \mathbf{N}_o' \mathbf{R}_{co}' \left(\mathbf{X}_p - \mathbf{X}_c \right) \tag{3}$$

Equation 1 can also be expressed in the format of the *Direct Linear Transformation* or DLT (Abdel-Aziz and Karara, 1971),

$$x_p = \left(a_1 X_p + a_2 Y_p + a_3 Z_p + a_4 \right) / D_p \tag{4a}$$

$$y_p = \left(a_5 X_p + a_6 Y_p + a_7 Z_p + a_8 \right) / D_p \tag{4b}$$

$$D_p = \left(a_9 X_p + a_{10} Y_p + a_{11} Z_p + 1 \right) \tag{4c}$$

In this form, the DLT is obtained after normalizing with respect to the projection of $\mathbf{X}_{co}$ onto the camera's optical axis. Thus, this factor should not vanish, which can be ensured by choosing the origin of E_o somewhere in the observed volume. For constant Z_p, Eq. 4 describes the general perspective transformation between two planes (X,Y) and (x,y).

It is of interest to note that rectilinearity is maintained under the perspective transformations of Eq. 1 and 3. Of course, these equations apply to the ideal case only; in practice, image distortion (see below) must be modeled. Furthermore, lens properties will result in defocusing errors unless the "depth-of-field" is relatively small, and the principal distance c will be slightly larger than the camera's focal length f, by virtue of the lens equation

$$1/c+1/\left|Z_{po}-Z_{co}\right|=1/f \tag{5}$$

for some average distance $|Z_{po} - Z_{co}|$ along the camera's optical axis.

In Röntgenographic systems, the perspective center is not formed by a lens, but by the Röntgen source, and the object under study is placed between the source and the image plane, usually a photographic plate. The same mathematical model applies, although distortion models may be different; since no lenses are used, nonlinear distortion is usually negligible.

3-D Point and Rigid-Body Reconstruction

In the case of general 3-D motion, single camera measurements generally do not yield sufficient information. In principle, combination of single-camera measurements and known body-fixed coordinates $\mathbf{X}_{bi}$ for at least three noncollinear markers P_i allow full recovery of a rigid body's position and attitude, but translations to and from the camera are very ill-determined from noisy image data (see Chapter 10.1). If the landmark distribution is planar and normal to the line between the landmark mean and the camera, also rotations normal to this line are ill-determined. It is therefore advisable to use multiple cameras for general 3-D motion studies.

Ideally, the observed direction lines when reconstructed from known (calibrated) camera stations should intersect at the true position of an observed point; because of measurement and calibration errors, this will usually not be the case, and some optimal estimate should be found. One such estimate is the ODLE-solution (Object Distance Lease-Squares Error), where the position $\mathbf{X}_{po_i}$ is selected which has the smallest rms distance to all reconstructed direction lines. In the general case, given camera stations $\mathbf{X}_{co_j}$ and observed normalized directions $\mathbf{N}_{po_{ij}}$ result in the following solution for $\mathbf{X}_{po_i}$:

$$\mathbf{X}_{po_i}=\left[\sum_{j=1}^{Nc}\left(I-\mathbf{N}_{po_{ij}}\cdot\mathbf{N}'_{po_{ij}}\right)\right]^{-1}\sum_{j=1}^{Nc}\left(I-\mathbf{N}_{po_{ij}}\cdot\mathbf{N}'_{po_{ij}}\right)\mathbf{X}_{co_j} \tag{6a}$$

where,

$$\mathbf{N}_{po_{ij}}=\mathbf{S}_{po_{ij}}/\left|\mathbf{S}_{po_{ij}}\right|,\qquad \mathbf{S}_{po_{ij}}=\mathbf{R}'_{co_j}\left(\mathbf{x}_{pij}-\mathbf{x}_{c_j}\right) \tag{6b}$$

which can be simplified to a vector product in the common case of $N_c = 2$ cameras (Woltring, 1980).

Alternatively, one may resort to the DLT-equations. After multiplying the denominator Di in Eq. 4 to the left and reordering, one finds a linear system of equations in $\mathbf{X}_{po_i}$. Combination of these equations for multiple cameras then results in the system

$$B_i \cdot \mathbf{X}_{po_i} = \mathbf{b}_i \tag{7}$$

where B_i and $\mathbf{b}_i$ are functions of the camera parameters and of the image coordinates, with least-squares solution

$$\mathbf{X}_{po_i} = (B_i' B_i)^{-1} B_i' \mathbf{b}_i \tag{8}$$

In many cases, the measurements $\{x_{p_{ij}}, y_{p_{ij}}\}$ can be regarded as affected by additive, zero-mean, uncorrelated, stochastic errors with known variances, possibly apart from a common scaling factor σ^2 (the "unit variance"). Because of the measurement errors in B_i, the estimator (Eq. 8) will generally yield biased estimates (cf. Section 10.1) and some form of least-squares, iterative adjustment calculus as required. In the case of the ODLE- and DLT-models, two or three iterations are usually sufficient.

From reconstructed landmark coordinates $\{\mathbf{X}_{po_i}\}$, the position and attitude of a rigid body with known body-fixed landmark coordinates $\{\mathbf{X}_{b_i}\}$ is the next step. Spoor and Veldpaus (1980) and Veldpaus et al. (1988) have described two algorithms for this purpose.

Given $m \geq 3$ landmark coordinate pairs $\{\mathbf{X}_{o_i}, \mathbf{X}_{b_i}\}$ in the object- and body-fixed coordinate systems E_o and E_b, the least-squares solution for the attitude matrix $\mathbf{R}_{bo}$ and for the position $\mathbf{X}_{bo}$ of the origin of E_b in E_o follow from the polar decomposition (Eq. 9a)

$$\mathbf{X}_o \cdot \mathbf{X}_b' = \mathbf{R}_{bo}' \cdot D, \qquad D = D' - \mathbf{R}_{bo}' \cdot \mathbf{X}_b \tag{9a}$$

$$\mathbf{X}_{bo} = \mathbf{X}_o - \mathbf{R}_{bo}' \cdot \mathbf{X}_b \tag{9b}$$

where $\mathbf{X}_o$ and $\mathbf{X}_b$ are the landmark means in the two coordinate systems and $\mathbf{X}_o$ and $\mathbf{X}_b$ the matrices

$$\mathbf{X}_o = \left[\mathbf{X}_{o_1} - \mathbf{X}_o, \mathbf{X}_{o_2} - \mathbf{X}_o, \cdots, \mathbf{X}_{o_m} - \mathbf{X}_o\right] \tag{10a}$$

and

$$\mathbf{X}_b = \left[\mathbf{X}_{b_1} - \mathbf{X}_b, \mathbf{X}_{b_2} - \mathbf{X}_b, \cdots, \mathbf{X}_{b_m} - \mathbf{X}_b\right] \tag{10b}$$

In the noise-free case, $D = \mathbf{X}_b \cdot \mathbf{X}_b'$; in the noisy case, the solution can be found by solving a certain cubic or quartic equation (Veldpaus et al., 1986).

If at least three landmark coordinate pairs are noncollinear in both coordinate systems, a unique solution can be found.

In the case of partial observations, e.g., landmark 1 is only seen by camera 1, landmark 2 by camera 2, and landmark 3 by both cameras, the above algorithms cannot be used. Here, some form of generalized, iterative adjustment calculus may be needed, based on linearized equations in terms of $\mathbf{X}_{bo}$, $\mathbf{R}_{bo}$, $\{\mathbf{X}_b\}$, the camera parameters, and the observed image coordinates (Woltring, 1982). Multiple solutions may occur, and the proper one should be chosen on the basis of continuity considerations.

In the models discussed above, it was assumed that all image coordinates and landmarks had been properly identified. If the identity of some data is unknown, one may use the rms fitting error to decide between candidate solutions. For example, one may calculate the predicted position of a landmark from a candidate solution for a rigid-body's position and attitude and evaluate the image distances to those image data which are available. The smallest rms error then yields the selection criterion. Of course, such a criterion does not guarantee that the proper choice has been made, and visual inspection of the results may remain advisable.

Image Distortion Models

Until now, the camera has been viewed as ideal, without any image distortion due to lens errors, film shrinkage, or electronic processing errors. Film shrinkage errors are mainly linear, while lens errors are largely radial, along the line towards the principal point. Some additional, tangential errors normal to the radial direction may also occur. A conventional model for film and lens distortion (e.g., Marzan and Karara, 1975; Woltring, 1980) is as follows:

For given, raw image coordinates $(x_p, y_p)'$ and principal point $(x_c, y_c)'$, linear distortion is modeled in terms of separate principal distances c_x and c_y per image axis, and of an image "shearing factor" α. This results in "rectified" image coordinates $(x_{rec}, y_{rec})'$,

$$x_{rec} = \left(x_p - x_c\right) / c_x \tag{11a}$$

$$y_{rec} = \alpha \cdot x_{rec} + \left(y_p - y_c\right) / c_y \tag{11b}$$

which are subsequently used in a lens-distortion model for calculating "refined" image coordinates $(x_{ref}, y_{ref})'$,

$$x_{ref} = b \cdot x_{rec} + p \cdot \left(r^2 + 2 \cdot x_{rec}^2\right) + 2 \cdot q \cdot x_{rec} \cdot y_{rec} \tag{12a}$$

$$y_{ref} = b \cdot y_{rec} + q \cdot \left(r^2 + 2 \cdot y_{rec}^2\right) + 2 \cdot p \cdot x_{rec} \cdot y_{rec} \tag{12b}$$

where,

$$r^2 = x_{rec}^2 + y_{rec}^2, \quad b = 1 + b_1 \cdot r^2 + b_2 \cdot r^4 + b_3 \cdot r^6 \tag{12c}$$

Radial distortion is modeled by the $\{b_i\}$, and tangential distortion by p and q.

Notwithstanding the validity of central projectivity, some investigators have reported satisfactory results with completely empirical models. As long as such models are used within the calibrated volume (cf. Wood and Marshall, 1986), this is certainly acceptable, especially if the model is numerically more efficient. For example, Fioretti et al. (1985) have described a highly accurate polynomial model which was twice as fast as the DLT; they motivated their model choice by pointing out that central projectivity is useful for ideal cameras, but that image distortion effects make a completely black-box approach more attractive if the observation volume is sufficiently small.

PHOTOGRAMMETRIC CALIBRATION METHODS

Traditional, photogrammetric camera calibration methods resort to distinct procedures for estimating the so-called internal and external camera parameters. The former (principal point and distance; image distortion parameters) are typically obtained via tests on an optical bench; the latter (camera position and attitude in object-space E_o) are obtained by surveying techniques or by observing a distribution of target markers with known ("absolute control") or unknown ("relative control") coordinates in E_o. Dapena et al. (1982) have described such a method in the context of sports biomechanics.

The current trend in photogrammetry is to simultaneously estimate all parameters from partial control in-the-field, given a suitable image distortion model. At the expense of more complicated software, simpler calibration structures can be used, and better guarantees exist that image distortion is properly modeled.

Direct Linear Transformation (DLT)

A simple calibration procedure is based on the Direct Linear Transformation which in its original form encompasses all forms of linear image distortion (Abdel-Aziz and Karara, 1971). For a given point P_i with coordinates $X_i = (X_i, Y_i, Z_i)'$ in E_o and image coordinates $(x_i, y_i)'$, Eq. 4 can be rearranged as:

$$\begin{Bmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -x_iX_i & -x_iY_i & -x_iZ_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -y_iX_i & -y_iY_i & -y_iZ_i \end{Bmatrix} \mathbf{a} = \begin{Bmatrix} x_i \\ y_i \end{Bmatrix} \tag{13}$$

When these equations are merged for a 3-D distribution of at least 6 targets P_i , a system similar to (7) is obtained, with solution as in (8). Again, biasedness is to be expected if the $(x_i, y_i)'$ are noisy, and iterative adjustment calculus is called for.

The 11 DLT-parameters can be identified with the conventional camera parameters of Eqs. 1 and 11. Combination of these equations and elimination of λ in Eq. 1 results in the collinearity equations of photogrammetry, generalized to accommodate all forms of linear distortion modeled in Eq. 11,

$$\begin{aligned} x_{rec} + \mathbf{r}'_{co_x}\left(\mathbf{X}_p - \mathbf{X}_{co}\right) / \mathbf{r}'_{co_z}\left(\mathbf{X}_p - \mathbf{X}_{co}\right) = 0 \\ y_{rec} + \mathbf{r}'_{co_y}\left(\mathbf{X}_p - \mathbf{X}_{co}\right) / \mathbf{r}'_{co_z}\left(\mathbf{X}_p - \mathbf{X}_{co}\right) = 0 \end{aligned} \tag{14}$$

where $\mathbf{R}_{co} = (\mathbf{r}_{co_x}, \mathbf{r}_{co_y}, \mathbf{r}_{co_z})'$ as in Eq. 1. Identification of corresponding elements in Eqs. 11, 13 and 14 yields the following relation:

$$\begin{Bmatrix} a_1 & a_2 & a_3 & a_4 \\ a_5 & a_6 & a_7 & a_8 \\ a_9 & a_{10} & a_{11} & 1 \end{Bmatrix} = \left(\mathbf{r}'_{co_z}\mathbf{X}_{co}\right)^{-1} \begin{Bmatrix} c_x & 0 & -x_c \\ -\alpha \cdot c_y & c_y & -y_c \\ 0 & 0 & -1 \end{Bmatrix} \left[\mathbf{R}_{co} \middle| -\mathbf{R}'_{co}\mathbf{X}_{co}\right] \tag{15}$$

For given conventional camera parameters, the DLT-parameters follow explicitly; for given DLTparameters, the conventional parameters follow via orthogonal triangularization and back-substitution (cf. Woltring, 1980, 1982).

Marzan and Karara (1975) have generalized the DLT to accommodate the non-linear distortion model (12); they have also included weight factors to account for inaccuracies in the control point and image coordinates. Their software package constitutes an elegant procedure for routine calibration in-the-field if a sufficiently large, 3-D absolute control distribution is available.

Analytical Selfcalibration

If the volume of interest is larger than, say, $1m^3$, the construction of a suitable calibration object for the DLT becomes prohibitive. Yet, the calibration volume should encompass the full volume lest extrapolation errors will assume significance, especially if non-linear image distortion is carried in the solution (cf. Wood andMarshall, 1986). For these reasons, the use of other photogrammetric procedures may be called for.

The most comprehensive of these is known as Analytical Selfcalibration (cf. Woltring, 1980). It appears possible to calibrate all parameters for one or more cameras by viewing a relative control point distribution (which may be planar) at different camera attitudes, and possibly from different

positions. It is not necessary to know the coordinates of all points or any of the camera parameters; it is merely required that the internal camera parameters be stable between exposures. A coordinate system must be defined, e.g., in terms of one camera exposure or of three known, non-collinear control points, and all remaining unknowns are estimated via iterative, linearized and weighted least-squares. The name "selfcalibration" derives from the fact that the unknown control point coordinates are usually the quantities of interest, but there is little against first estimating the camera parameters, to be used as known constants in subsequent reconstruction of observed movement. In this fashion, a very simple calibration distribution without high accuracy can be used.

Simultaneous Multiframe Analytical Calibration (SMAC)

The implementation of Analytical Selfcalibration as a software package is quite complicated, and various investigators have looked into intermediate solutions. Woltring (1980) has described a method where a known, planar distribution of control points is used. Such a planar grid can be easily constructed, transported and manipulated if the volume of interest is not larger than a few m^3, e.g., as in gait analysis. The planar nature of the grid renders it unsuitable for the DLT, but it has some of the advantages of Analytical Selfcalibration.

It is sufficient to position the cameras in a convergent way, i.e., the optical axes are not parallel, and to observe the grid while it is held in a number of oblique attitudes with respect to the cameras: see Fig. 4. Except for one defined reference position and attitude, the positions and attitudes of the grid are unknowns together with the unknown camera parameters. The calibrated volume can be enlarged by positioning the grid at other positions, as long as it is observed by at least two convergent cameras. In a circular measurement configuration, the control points may be chosen on both sides of the planar grid.

This "Simultaneous Multiframe Analytical Calibration" procedure (SMAC) allows to cover larger volumes than is practically feasible with the DLT. For very large volumes, e.g., in competitive sports, Analytical Selfcalibration may be more suitable.

The calibration and reconstruction package of Selvik (1974) can be run in a DLT-like mode and in a SMAC-like mode. The package assumes that the calibration points have accurately known positions in two parallel planes with known distance between the planes: see Fig. 5. However, one rotational and two positional degrees of freedom of these planes with respect to each other may entail inaccuracies. By viewing the object in one position, and once again after a rotation of approximately 180 degrees about the normal on the two planes, the inaccuracies can be assessed and eliminated.

This feature renders the construction of highly accurate, 3-D calibration objects relatively easy.

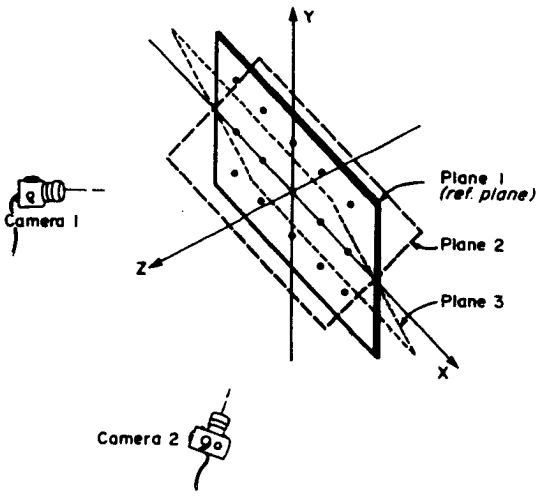

Fig. 4

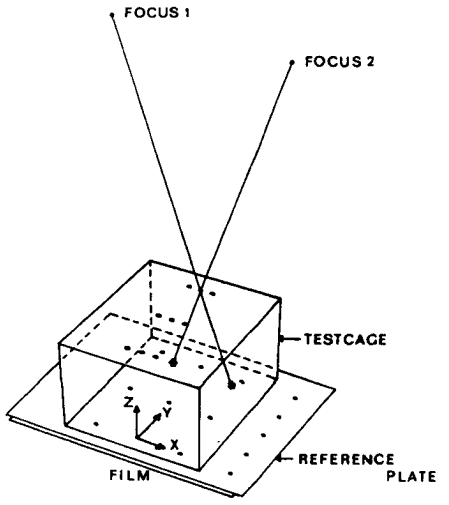

Fig. 5

Figure 4. Calibration grid and camera lay-out in SMAC (from Woltring, 1980)

Figure 5. Schematic representation of a Röntgenstereophotogrammetric calibration exposure. The object-space co-ordinate system is fixed with respect to the lower plane of the test-cage, similar to the "reference plane" in the SMAC-method. After the calibration exposure, the "reference plane" serves to define the position and attitude of the photographic plate with respect to object-space (after van Dijk, 1983)

FUTURE DEVELOPMENTS

The most important tendency is that movement analysis by photogrammetry will become more "intelligent". If there is no loss of data due to shadowing effects or other artifacts, the calibration and reconstruction procedures can be of a straightforward algorithmic nature, with little need to adjust the computational process. However, many aspects of practical movement analysis exhibit measurement artifacts and more complex signal properties than can be modeled at present, thus necessitating trial-and-error interventions. Obvious examples are the identification of passive markers in time and between cameras, especially when a marker reappears after temporary obscurity, and the interpolation of data gaps during obscurity. Also, skin movement artifacts should be addressed in this way.

The presence of well-identifiable landmarks in multiple images has been the primary assumption of this chapter. Artificial markers may be easily applied in a research context, both on the skin and as implanted metallic

markers in bone material, but they have limited applicability in a routine clinical and sports-training environment, while natural landmarks provide insufficient information for reliable 3-D kinematics assessment. For example, implanted markers have been used by de Lange et al. (1986) for in vitro study of the kinematics of individual carpal bones, and Benink (1985) used a similar approach for the ankle joint. There is a high need for non-invasive methods to study the internal kinematics of such complex joints.

Current developments in pattern recognition and artificial intelligence (e.g., Young and Fu, 1986) hold promise for 3-D motion reconstruction based on shape analysis of asymmetrical, rigid bodies. For example, one could obtain a 3-D shape description of individual carpal or ankle bones by means of Magnetic Resonance Imaging. This information could subsequently be related to projected Röntgen images of the joint.during motion, as observed via low-intensity Röntgen sources and image-intensifying radiation detectors. Identification of observed contours will allow 3-D movement reconstruction of the individual bones if their 3-D shapes are sufficiently asymmetrical; this is certainly the case in complex joints like the wrist and ankle.

Microcomputer-controlled systems and digital signal processing IC's are appearing on the market, thus allowing cheaper and real-time data processing. While most data processing is currently conducted via a sequence of logically distinct subprocedures (independent marker identification per camera, landmark position reconstruction, smoothing, and interpolation, followed by rigid-body calculus and finite differencing for derivative estimation), the field of signal processing allows the use of more general and universal data conditioning procedures where even incomplete information is optimally exploited. Similar developments can be observed in robotics, scene analysis, spatial navigation, and process control (e.g., Gelb, 1974).

The availability of faster (and less expensive!) hardware and software in biokinematics should result in increased utilization of movement analysis in the research and clinical fields. Up to now, most work in movement analysis has been either of a purely kinematic nature with emphasis on simple, directly measurable quantities in large numbers of patients (e.g., angle-angle diagrams), or of a largely research nature with further modeling and data processing on rather small numbers of healthy subjects or patients. Combination of biokinematic quantities with geometric and anatomical data as discussed elsewhere in this book should render movement analysis by photogrammetry a practical tool in many situations, both in vitro and in vivo, in particular for assessing kinetic (loading) situations in the musculoskeletal system.

ACKNOWLEDGMENT

The authors are indebted to Philips Medical Systems Nederland B.V. for the facilities enjoyed while preparing this paper.

REFERENCES

Abdel-Aziz, Y. I. and Karara, H. M. (1971) "Direct linear transformation from comparator coordinates into object-space coordinates," in *Proceedings ASP/UI Symposium on Close-Range Photogrammetry,* pp. 1-18, American Society of Photogrammetry, Falls Church, VA.

Benink, R. J. (1985) "The constraint-mechanism of the human tarsus a Ro ntgenological experimental study," *Acta Orthop. Scand.*, Vol. 56, Supplement 215.

Brumbaugh, R. B., Crowninshield, R. D., Blair, W. F. and Andrews, J. G. (1982) "An in vivo study of normal wrist kinematics," *J. Biomech.* Eng., Vol. 104, pp. 176-181.

Dapena, J., Hartman, E. A. and Miller, E. A. (1982) "Three-dimensional cinematography with control object of unknown shape," *J. Biomech.*, Vol. 15(1) pp. 11-19.

Dijk, R. van (1983) "The behavior of the cruciate ligaments in the knee," Dissertation, University of Nijmegen, The Netherlands.

Ferrigno, G. and Pedotti, A. (1985) "ELITE - a digital dedicated hardware system for movement analysis via real-time TV signal processing," *IEEE Trans. Biomed. Eng.*, Vol. BME 32(11) pp. 943-950.

Fioretti, S., Germani, A. and Leo, T. (1985) "Stereometry in very close-range stereophotogrammetry with nonmetric cameras for human movement analysis," *J. Biomech.*, Vol. 18(11) pp. 831-842.

Furnée, E. H. (1986) "Innovation in video-digital coordinate-measurement for movement analysis," These Proceedings.

Gelb, Ed., A. (1974) *Applied Optimal Estimation*, The M.I.T. Press, Cambridge, MA.

Grieve, D. W., et. al (1975) *Techniques for the Analysis.of Human Movement*, Lepus Books, London.

Güth, V., Abbink, F. and Heinrichs, W. (1973) "Eine Method zur chronozyklographischen Bewegungsaufzeichnung mit einem Prozeszrechner," *Int. Z. Angew. Physiol.*, Vol. 31, pp. 151-162.

Heinrichs, W. (1974) "Eine digitale Zeitmesz-Einrichtung hoher Auflösung zur Weiterentwicklung der chronocyclographischen

Bewegungsaufnahme mit Prozeszrechner," *Europ. J. Appl. Physiol.*, Vol. 32, pp. 227-238.

Huiskes, R., Kremers, J., Lange, A. de, Woltring, H. J., Selvik, G. and Rens, Th. J. G. van (1985) "Analytical stereophotogrammetric determination of three-dimensional knee-joint geometry," *J. Biomech.*, Vol. 18(8) pp. 559-570.

Lange, A. de, Kauer, J. M. G., Huiskes, R. and Woltring, H. J. (1986) "Carpal bone motion axes and pivots in flexion and extension of the hand," in *Proceedings 32nd Annual Orthopaedic Research Society Meeting*, New Orleans, LA.

Lindholm, L. E. and Öberg, K. E. (1974) "An opto-electronic instrument for remote on-line movement monitoring," *Biotelemetry*, Vol. 1(2) pp. 94-95.

Macellari, V. (1983) "CoSTEL a computer peripheral remote sensing device for 3-dimensional monitoring of human motion," *Med. Biol. Eng. Comp.*, Vol. 21(3) pp. 311-318.

Marzan, G. T. and Karara, H. M. (1975) "A computer program for direct linear transformation solution of the collinearity condition and some applications for it," in *Proceedings Symposium on Close-Range Photogrammetric Systems*, pp. 558-574, American Society of Photogrammetry, Falls Church, VA.

Mesqui, F., Kaeser, F. and Fischer, P. (1985) "Real-time noninvasive recording and three-dimensional display of the functional movements of an arbitrary mandible point," in *Biostereometrics -'85*, ed. A. M. Coblentz and R. E. Herron.

Selvik, G. (1974) *A Röntgen stereophotogrammetric method for the nonlinear least-squares problem*, Ph.D. Thesis, AvCentralen, Lund, Sweden.

Selvik, G. (1983) "Röntgen stereophotogrammetry in orthopaedics," in Biostereometrics '82, ed. R. E. Herron and A. M. Coblentz, pp. 178-185.

Slama, J.J., Ed. (1980) *Manual of Photogrammetry*, American Society of Photogrammetry, Falls Church, VA.

Spoor, C. W. and Veldpaus, F. E. (1980) "Rigid-body motion calculated from spatial coordinates of markers," *J. Biomech.*, Vol. 13(4) pp. 391-393.

Taylor, K. M., Mottier, F. M., Simmons, D. W., Cohen, W., Pavlak, R. Jr., Cornell, D. P. and Hawkins, G. B. (1982) "An automated

motion measurement system for clinical gait analysis," *J. Biomech.*, Vol. 15(7) pp. 505-516.

Veldpaus, F. E., Woltring, H. J. and Dortmans, L. J. G. M. (1988) "An algorithm for the equiform transformation from spatial marker coordinates," *J. Biomech.*, Vol. 21(1) pp. 45-54.

Winter, D. A., Greenlaw, R. K. and Hobson, D. A. (1972) "Television computer analysis of kinematics of human gait," *Comp. Biomed. Res.*, Vol. 5, pp. 498-504.

Woltring, H. J. (1975) "Single- and dual-axis, lateral photodetectors of rectangular shape," *IEEE Trans. Electron. Devices*, Vol. 22(8) pp. 581-590.

Woltring, H. J. (1980) "Planar control in multi-camera calibration for three-dimensional gait studies," *J. Biomech.*, Vol. 13(1) pp. 39-48.

Woltring, H. J. (1982) "Estimation and precision of three-dimensional kinematics by analytical photogrammetry," in *Computing in Medicine*, ed. J. P. Paul, et al., pp. 232-241, The MacMillan Press, London/Basingstoke.

Woltring, H. J. (1984a) "On methodology in the study of human movement," in *Human Motor Actions - Bernstein Reassessed*, pp. 35-73, North-Holland, Amsterdam.

Woltring, H. J. (1984b) "Locomotion assessment with infrared and visible light," in *Medical Telemetry II*, ed. H. P. Kimmich and M. Bornhausen, pp. 95-106.

Woltring, H. J., Huiskes, R. and Lange, A. de (1985) "Finite centroid and helical axis estimation from noisy landmark measurements in the study of human joint kinematics," *J. Biomech.*, Vol. 18(5) pp. 379-389.

Woltring, H. J. and Marsolais, E. B. (1980) "Opto-electronic (SELSPOT) gait measurement in two- and three-dimensional space — a preliminary report," *Bull. Prosthetics Research*, pp. 46-52.

Wood, G. A. and Marshall, R. N. (1986) "The accuracy of DLT extrapolation in three-dimensional film analysis," *J. Biomech.*, Vol. 19(9) pp. 781-785.

Young, T. Y. and Fu, K. S. (ed.) (1986) *Handbook of Pattern Recognition and Image Analysis*, Academic Press, Orlando, FL.