The International Journal of Robotics Research
http://ijr.sagepub.com

Autonomous Robot Calibration for Hand-Eye Coordination

David J. Bennett, Davi Geiger and John M. Hollerbach
The International Journal of Robotics Research 1991; 10; 550
DOI: 10.1177/027836499101000510

The online version of this article can be found at:
http://ijr.sagepub.com/cgi/content/abstract/10/5/550

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
Multimedia Archives

Additional services and information for *The International Journal of Robotics Research* can be found at:

Email Alerts: http://ijr.sagepub.com/cgi/alerts

Subscriptions: http://ijr.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

David J. Bennett
Davi Geiger
Massachusetts Institute of Technology
Artificial Intelligence Laboratory
Cambridge, Massachusetts 02139

John M. Hollerbach
McGill University
Department of Biomedical Engineering
Montreal, PQ, Canada, H3A 2B4

Autonomous Robot Calibration for Hand-Eye Coordination

Abstract

Autonomous robot calibration is defined as the process of determining a robot's model by using only its internal sensors. It is shown that autonomous calibration of a manipulator and stereo camera system is possible. The proposed autonomous calibration algorithm may obtain the manipulator kinematic parameters, external kinematic camera parameters, and internal camera parameters. To do this, only joint angle readings and camera image plane data are used. A condition for the identifiability of the manipulator/camera parameters is derived. The method is a generalization of a recently developed scheme for self-calibrating a manipulator by forming it into a mobile closed-loop kinematic chain.

1. Introduction

Hand-eye coordination is a particularly demanding task, because it requires the consistency of two separate sensing systems: the robot manipulator and the stereo vision system. It is the intention of this article to address the issue of how these two systems may be calibrated autonomously. "Calibration" means the determination of all of the parameters of the internal models of both the camera and arm systems. For the purposes of this article *autonomous calibration* is defined as an automated process that determines the model parameters by using only the robot's *internal sensors*. Thus for autonomous kinematic calibration, only the joint angle and 2D camera image sensors are permitted.

The International Journal of Robotics Research,
Vol. 10, No. 5, October 1991,
© 1991 Massachusetts Institute of Technology.

1.1. Motivation

Autonomy is an important property for a robot that must function outside of controlled laboratory conditions. It is inevitable that a robot will have its base moved, links bent, or cameras misaligned or be otherwise damaged. In such situations it would be desirable not to have to resort to the use of special-purpose calibration equipment to update the model used for robot control. In fact, an ultimate goal would be for the robot to be able to calibrate its internal model in real time.

Although in certain engineering applications the goal of autonomy may be sacrificed in favor of simplicity, it is pointed out that humans have no such choice in calibrating their sensory-motor system. Thus a second motivation for studying autonomous calibration derives from an interest in understanding the human sensory-motor system (i.e., human motor control research) (Bennett 1990).

1.2. Previous Research

In the domain of robot *dynamics*, autonomous calibration has essentially been achieved, although the kinematics must be assumed to be known (An et al. 1988). In particular, it is possible to estimate the inertial parameters that define the various links by using only internal joint torque (current), position, and velocity measurements. This idea has actually been made to operate in real-time model-based adaptive control schemes (Niemeyer and Slotine 1988). The success of inertial estimation is based on the observation that the suitable combinations of the inertial parameters enter linearly into the dynamic equations.

In contrast, autonomously determining the *static*

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

relationship between the internal joint angle sensors and the manipulator end-point position—referred to as the *kinematic model*—has not been as successful as autonomous dynamic estimation (Hollerbach 1989). Typically, researchers have viewed the manipulator as a positioning device—that is, an open-loop kinematic chain. This view demands that the end point be measured in addition to the joint angles to infer the kinematic parameters. Therefore autonomous calibration is not possible.

If instead the manipulator is viewed as a device to interact with the environment, then autonomous calibration is possible (Bennett and Hollerbach, 1988; 1989a,b). The basic observation is that the manipulator may form a *mobile closed-loop kinematic chain* when performing a task. The internal model structure, the knowledge of the type of task constraint, and the internal joint angle measurements collected while in a number of configurations provide enough consistency equations to solve for the kinematic parameters.

For example, imagine a 6-DOF manipulator making a fixed-point contact (a 3-DOF task) with a surface (i.e., similar to holding a pen tip at a fixed location). The resulting closed-loop kinematic chain is mobile; so, if the closed loop is moved into n configurations, then the joint angle readings may be used to write $3n$ loop position equations that may be solved to identify all of the kinematic parameters (Bennett and Hollerbach 1989b). The transformations from robot base to the contact point and from the wrist to the contact point are also determined in the process. This technique is quite general. As an example that is most relevant to the hand-eye calibration problem, we point out that two uncalibrated manipulators may calibrate one another by moving while rigidly grasping a common object (a 0-DOF task) (Bennett and Hollerbach 1988).

Three observations are worth stressing:

- The knowledge of the *task constraint* (e.g., two robots gripping together) replaces the need for an external sensor.
- The *redundancy* of the sensing systems (e.g., two arms) with respect to the task enables the various redundant components to move while performing the same task.
- The *a priori model structure knowledge* allows one to form a number of consistency equations (the kinematics) that may be solved for the kinematic parameters.

These three observations serve as a basis for extending autonomous manipulator calibration to complete hand-eye calibration. First we review relevant vision system calibration techniques.

The conventional methodology for camera calibration is to move a number of known precision points into the field of view of the cameras and infer the camera calibration from the given points in space (Tsai 1989).

One effective approach is to form a look-up table from known rays (obtained from two planes of points in space) to recorded image locations (Martins et al. 1981; Gremban et al. 1988) and then use splines to do local interpolation. Look-up table approaches need external calibration points, and thus they must be disregarded for autonomous camera calibration.

Various model-based approaches have been used for camera calibration. In computer vision and graphics the pinhole camera model has been used extensively (Sobel 1974; Duda and Hart 1973; Yakimovsky and Cunningham 1978; Horn 1986). This model may be augmented to account for lens distortions (Moffitt and Mikhail 1980; Ziemann 1985; Tsai 1986). Because the pinhole camera model is nonlinear in its parameters, there have been various proposals to make the calibration equations linear (Abdel-Aziz and Karara 1971; Ganapathy 1984; Tsai 1986). These methods are important, because they provide initial guesses at the parameters that general nonlinear parameter search methods require. Other empirical polynomial interpolation models have also been used (Martins et al. 1981; Isaguirre et al. 1985; Gremban et al. 1988), their only advantage being that the parameters enter linearly into the equations.

These camera modeling techniques and parameter identification methods are relevant to autonomous calibration, but most of this work is predicated on the assumption that there are external calibration points available. There are a few notable exceptions. For example, early photogrammetric engineering work (Kenefick et al. 1972) and more recent robotics research (Brooks et al. 1988) have demonstrated that the camera parameters may be recovered by moving the cameras while viewing arbitrary unknown points in space.

Finally, there is a considerable body of literature that addresses the problem of registering the calibrated vision system's coordinates with respect to the robot base (Hollerbach 1989). Especially interesting is the work in Puskorius and Feldkamp (1987). Their technique is to determine the camera to hand transformation simultaneously to the robot parameters by viewing an arbitrary fixed point in space. The internal camera parameters are calibrated beforehand by viewing a precision calibration jig.

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

1.3. Toward Autonomous Hand-Eye Calibration

As can be seen, the calibration of a robot manipulator/vision system is typically based on a "divide and conquer" principle. None of these approaches may be made autonomous. Further, there is no guarantee that once the various separately calibrated components are assembled they will be *consistent*. This is an important point, as *models and sensors are usually inaccurate*. We thus propose that a solution is to calibrate the models of the manipulator and two cameras simultaneously while performing the *task* of interest—hand-eye coordination. This may be done as follows.

Recalling the three observations made concerning closed-loop manipulator calibration, it is remarked that hand-eye calibration fits into this framework. First, a manipulator and a vision system may sense the location of the same point in space, and thus the total robot sensing is *redundant*. Second, if the *task* is defined as the cameras tracking the hand, then a closed kinematic loop is formed. This task constraint replaces the need for external calibration points. Finally, because we may assume a priori knowledge of both the *camera and the arm kinematic model structure*, it is possible to write consistency equations of the closed loop in a number of configurations and thus solve for the parameters. We will develop this idea in the following.

1.4. Outline

As stated, the purpose of this article is to extend the "closed-loop" approach developed for calibrating robot arms to calibrating a complete robot—with a stereo vision system in addition to the manipulator. The stereo system is assumed to have one axis of rotation per camera but is otherwise taken to have an arbitrary geometry.

An uncalibrated stereo camera system will be made to track a point on the hand of an uncalibrated arm. There are at least two distinct approaches to forming the closed-loop calibration equations. The first is to formulate a model for the manipulator relative to each camera separately and measure the position error in 2D image plane coordinates. The calibration would proceed by collecting data from manipulator/camera movements and minimizing the *image plane* error in both cameras. The second approach is to directly model the position of the end effector given by the stereo calculation (from the image data). The calibration can then proceed by minimizing the *end-effector error* between the manipulator and stereo models. Of these two approaches, the second is chosen, because it seems more natural to minimize the *task space* error. In addition, using the second approach enables one to formulate the iterative identification equations more simply; in particular, the manipulator Jacobians developed in Bennett and Hollerbach (1988) are directly applicable.

It is assumed that the point that is to be tracked may be unambiguously located on both camera images. This generally nontrivial correspondence problem (Horn 1986) may be solved here because of two additional constraints. First, it is known that the hand is moving relative to the background; therefore it is possible to disambiguate the hand image from the background. Second, a convenient point that can always be unambiguously located on the hand may be used (e.g., the tip of a pointer).

2. Model Definitions

2.1. Manipulator Model

Consider an arbitrary manipulator with n_f degrees of freedom. Let the 4 × 4 homogeneous transformation A_j from link j to link $(j - 1)$ be defined by the Denavit-Hartenberg (D-H) convention (Denavit and Hartenberg 1955) given in Figure 1 and symbolically as:

$$A_j = Rot(z, \theta_j')Trans(z, s_j)Trans(x, a_j)Rot(x, \alpha_j)$$

where the notation $Rot(x, \phi)$ indicates a rotation about an axis x by ϕ and $Trans(x, a)$ indicates a translation along an axis x by a. The joint angle is presumed to be related to the sensor reading by a constant offset: $\theta_j' = \theta_j + \theta_j^{off}$.

For convenience we define the base of the manipulator to coincide with a head-referenced coordinate system that is coincident with the left camera's axis of rotation (see next section). The position of the last link is related to these base coordinates by a sequence of D-H transformations defining the kine-

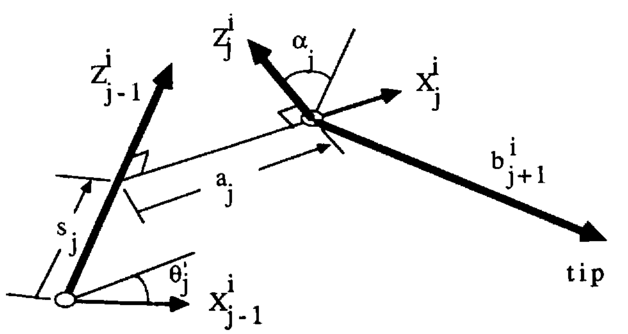

Fig. 1. Denavit-Hartenberg coordinates and tip vector $\mathbf{b}_j^i$.

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

matic model:

$$T_c = A_0 A_1 A_2 \ldots A_{nf}. \tag{1}$$

Because the model must only locate a point on the end effector, the last axis skew is unnecessary:

$$\alpha_{nf} = 0.$$

The position of the point P to be tracked is thus given by the *minimal parameter model:*

$$\begin{bmatrix} \hat{\underline{x}} \\ 1 \end{bmatrix} = A_0 A_1 A_2 \ldots A_{nj}\underline{o} \tag{2}$$

where the "^" emphasizes that it is the position of P *modeled* by the manipulator system, and $\underline{o} = (0, 0, 0, 1)$ is the location of the point P in the hand-based reference frame.

All of the *unknown* manipulator kinematic parameters are placed into an array:

$$\hat{\underline{\phi}} = (\underline{\theta}^T_{off}, \underline{s}^T, \underline{a}^T, \underline{\alpha}^T)^T$$

where $\underline{s} = (s_1, s_2, \ldots)^T$, etc.

2.2. *Visual System Model*

The model of a stereo camera system may be decoupled into a purely geometric part giving the relative orientation and position of the cameras and a part modeling each camera's projection of points in space. The parameters for these two portions are respectively called the *external* and *internal camera parameters*.

The internal camera model used is the standard pinhole camera model (Haralick 1980; Horn 1986). More refined parametric models including lens distortions (Moffitt and Mikhail 1980; Ziemann 1985; Tsai 1986) may also be incorporated without changing the general approach. Let the effective focal point distance be denoted by f and the projected point P in the image plane be given by the pair (u, v). Further define local camera coordinates to have x- and y-axes parallel to the camera (u, v) grid and have an origin at the focal point (Fig. 2). Thus the coordinates (x_R, y_R, z_R) of a point P expressed in the right camera's local coordinate system are given by the standard projection equations:

$$z_R/(-f_R) = x_R/(u_R + u_R^{off}) = y_R/(v_R + v_R^{off}). \tag{3}$$

Notice that provision is included for unknown offsets (u_R^{off}, v_R^{off}) between the image plane readings and the optical center of the camera. Analogously, the left-camera equations are:

$$z_L/(-f_L) = x_L/(u_L + u_L^{off}) = y_L/(v_L + v_L^{off}). \tag{4}$$

The *external* geometric model of the two camera system can be represented by D-H transformations. To distinguish the camera D-H parameters, a tilde (e.g., $\tilde{s}_0$) is used. We assume that each camera has one degree of rotation about a fixed axis with a joint angle sensor ($\theta_L = \tilde{\theta}_2$ and $\theta_R = \tilde{\theta}_1$ for the left and right cameras, respectively) as defined below. It is convenient to start the kinematic chain at the right camera's local coordinate system—that is, the frame located at the right camera's focal point and having its x- and y-axes parallel to the right image plane (u, v) coordinate grid. As mentioned in the previous section, the "base" coordinates are assumed to be located at the left camera axis of rotation. Thus the transformation of a point $\underline{x} = (x, y, z, 1)^T$ in base coordinates to the local coordinates of the right camera $\underline{x}_R = (x_R, y_R, z_R, 1)^T$ is given by (Fig. 3):

$$\underline{x}_R = \tilde{A}_0 \tilde{A}_1 \underline{x} = T_R^{-1}\underline{x}, \tag{5}$$

where the only variable parameter is $\tilde{\theta}'_1 = \theta_R + \tilde{\theta}_1^{off}$, the right camera rotation. We also define the opposite transform as $\underline{x} = T_R \underline{x}_R$. The left camera

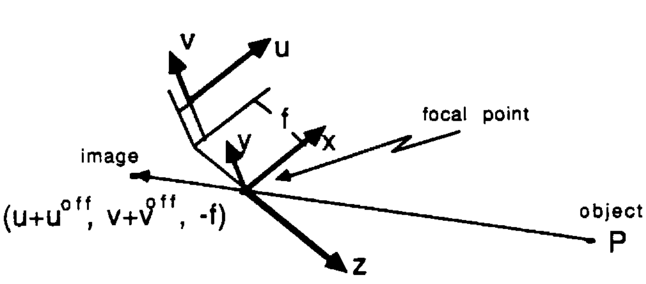

Fig. 2. The internal camera model.

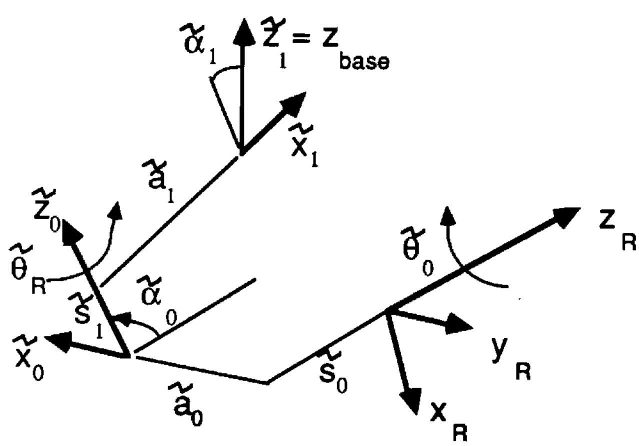

Fig. 3. The right camera axes to base D-H transformations. The first transformation locates the line of action of the right camera rotation. The second transformation locates the base coordinate z-axis, which is also the left camera axis of rotation.

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

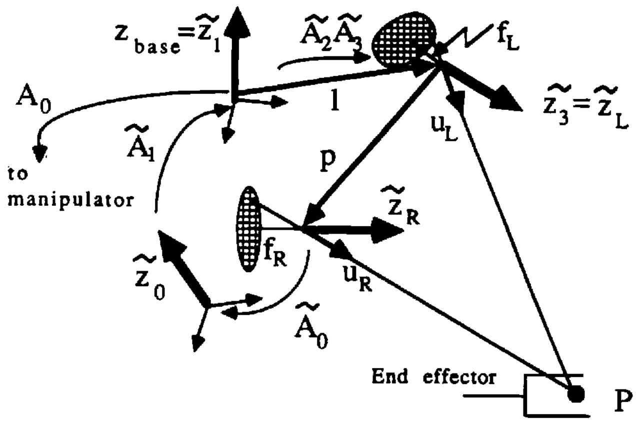

Fig. 4. A stereo camera system attached to a manipulator. L and R indicate the left and right cameras. The left and right cameras rotate about $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_0$, respectively.

coordinate system may be located by two additional D-H transformations (Fig. 4):

$$\underline{x} = \tilde{A}_2\tilde{A}_3\underline{x}_L = T_L\underline{x}_L \tag{6}$$

where the only variable parameter is $\tilde{\theta}_2' = \theta_L + \tilde{\theta}_2^{off}$, the left camera rotation. The parameters $\tilde{s}_3$ and $\tilde{\theta}_3$ translate and orient about the left optical axis so that the left coordinate system lines up with the (u, v) image grid and is located at the focal point. The parameters $\tilde{a}_3$ and $\tilde{\alpha}_3$ are redundant and are taken to be fixed zero quantities. We also define $\boldsymbol{R}_R$ and $\boldsymbol{R}_L$ as the upper 3×3 rotation matrices of $\boldsymbol{T}_R$ and $\boldsymbol{T}_L$, respectively.

To solve the stereo camera equations, it is convenient to define the vector $\mathbf{p}$ from the left focal point to the right focal point, the vector $\mathbf{l}$ from the base frame origin to the left focal point, and the two internal parameter vectors as follows:

$$\mathbf{p} = -\boldsymbol{R}_L\mathbf{p}_L, \qquad \begin{bmatrix} p_L \\ 1 \end{bmatrix} = \tilde{A}_0\tilde{A}_1\tilde{A}_2\tilde{A}_3\underline{o} \tag{7}$$

$$\begin{bmatrix} \mathbf{l} \\ 1 \end{bmatrix} = \boldsymbol{T}_L\underline{o} \tag{8}$$

$$\mathbf{u}_L = -\boldsymbol{R}_L \begin{bmatrix} u_L + u_L^{off} \\ v_L + v_L^{off} \\ -f_L \end{bmatrix} \tag{9}$$

$$\mathbf{u}_R = -\boldsymbol{R}_R \begin{bmatrix} u_R + u_R^{off} \\ v_R + v_R^{off} \\ -f_R \end{bmatrix} \tag{10}$$

where again $\underline{o} = (0, 0, 0, 1)^T$. Notice that $\mathbf{u}_L$ and $\mathbf{u}_R$ are vectors from the left and right focal points, respectively, along the line of sight. Thus the point P in base coordinates is simply:

$$\tilde{\underline{x}} = c\mathbf{u}_L + \mathbf{l}, \tag{11}$$

where the "$\tilde{\ }$" emphasizes that it is the position of P modeled by the camera system. The scalar c is given by:

$$c = \frac{\mathbf{u}_R \times \mathbf{u}_L \cdot \mathbf{u}_R \times \mathbf{p}}{\mathbf{u}_R \times \mathbf{u}_L \cdot \mathbf{u}_R \times \mathbf{u}_L} \equiv \Gamma(\mathbf{p}) \tag{12}$$

where "$\times$" denotes vector cross product, and "$\cdot$" denotes inner product. The *linear* operator $\Gamma(\cdot)$ has been defined here, as it will be useful later.

All of the *unknown* camera parameters are placed into an array:

$$\tilde{\underline{\phi}} = (\tilde{\underline{\theta}}_{off}^T, \tilde{\underline{s}}^T, \tilde{\underline{a}}^T, \tilde{\underline{\alpha}}^T, \underline{i}^T)^T$$

where $\underline{i} = (f_R, f_L, u_R^{off}, v_R^{off}, u_L^{off}, v_L^{off})^T$ and $\tilde{\underline{s}} = (\tilde{s}_1, \tilde{s}_2, \ldots)^T$, etc.

2.3. The Closed-Loop Model

Comparing equations (11) and (12), it is seen that the difference defines the closed-loop kinematic equations of the hand-eye system:

$$0 = \tilde{\underline{x}}(\tilde{\underline{\phi}}) - \hat{\underline{x}}(\hat{\underline{\phi}}) = \Delta\underline{x}(\underline{\phi}), \tag{13}$$

where we have defined $\underline{\phi} = (\hat{\underline{\phi}}^T, \tilde{\underline{\phi}}^T)^T$ as a concatenation of all of the parameters to be identified. Note that only *unknown* parameters need be included in $\underline{\phi}$. For example, if all the camera parameters are known and not included in $\underline{\phi}$, then the calibration method described would be a standard manipulator calibration scheme, as, for example, in Bennett and Hollerbach (1988).

Also, recall that the base coordinates of the manipulator were defined to correspond to the local coordinates of the left camera rotation. Thus we see that the manipulator base coordinates are the camera coordinates with axes $\tilde{x}_1$, $\tilde{y}_1$, and $\tilde{z}_1$.

3. Model Identification Procedure

As the cameras track the point P at discrete locations, the joint angle and image plane sensory information are recorded. For convenience the data recorded at the ith configuration are placed into a single array:

$$\underline{u}^i = (\theta_1^i, \ldots \theta_{n_f}^i, \theta_L^i, \theta_R^i, u_L^i, v_L^i, u_R^i, v_R^i)^T \tag{14}$$

At the ith configuration of the hand, one vector equation of the form (13) may be written

$$\underline{0} = \Delta\underline{x}^i = \Delta\underline{x}(\underline{u}^i, \underline{\phi}), \tag{15}$$

where in addition to the functional dependence on

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

the model parameters the dependence on the ith data array $\underline{u}^i$ is explicitly shown. As a short form, $\Delta \underline{x}^i$ will be used, where the functional dependence on the ith data array $\underline{u}^i$ will be understood by the superscript i. After moving the hand and cameras into m distinct locations, $3m$ scalar equations are generated. These equations must be solved in order to find the optimal parameter set $\underline{\phi}$, completing the calibration.

3.1. *Iterative Identification Technique*

Unfortunately, the equations of the kinematic loop are nonlinear, so we must use an iterative method to search for the solution $\underline{\phi}$. By expanding (13) with a Taylor's series about an initial guess $\underline{\phi}_0$ and neglecting higher order terms, the following linearized form is obtained:

$$\Delta \underline{x}^i = \left[\frac{\partial \underline{\bar{x}}^i(\underline{\phi}_0)}{\partial \underline{\phi}} - \frac{\partial \underline{\hat{x}}^i(\underline{\phi}_0)}{\partial \underline{\phi}} \right] \Delta \underline{\phi} = C^i \Delta \underline{\phi}, \quad (16)$$

where $\Delta \underline{\phi} = \underline{\phi}_0 - \underline{\phi}$. The Jacobian C^i may also be written as:

$$\left[\frac{\partial \hat{x}^i}{\partial \boldsymbol{\theta}} \,\middle|\, \frac{\partial \hat{x}^i}{\partial s} \,\middle|\, \frac{\partial \hat{x}^i}{\partial \boldsymbol{a}} \,\middle|\, \frac{\partial \hat{x}^i}{\partial \boldsymbol{\alpha}} \,\middle|\, \frac{\partial \bar{x}^i}{\partial \bar{\boldsymbol{\theta}}} \,\middle|\, \frac{\partial \bar{x}^i}{\partial \bar{s}} \,\middle|\, \frac{\partial \bar{x}^i}{\partial \bar{a}} \,\middle|\, \frac{\partial \bar{x}^i}{\partial \bar{\alpha}} \,\middle|\, \frac{\partial \bar{x}^i}{\partial i} \right].$$

Recall that $\Delta \underline{x}^i$ is the difference between the location of the point P given by the camera model and the manipulator model (computed using ϕ_0). Until the system is calibrated, this difference is nonzero.

Place the equations for each of the loop configurations into one array:

$$\Delta \underline{X} = \begin{bmatrix} \Delta \underline{x}^1 \\ \Delta \underline{x}^2 \\ \vdots \\ \Delta \underline{x}^m \end{bmatrix} = \begin{bmatrix} C_1 \\ C^2 \\ \vdots \\ C^m \end{bmatrix} \Delta \underline{\phi} = C \Delta \underline{\phi} \quad (17)$$

The solution that minimizes the modified-least-squares criteria $(\Delta \underline{X} - C\Delta \underline{\phi})^T(\Delta \underline{X} - C\Delta \underline{\phi}) + \lambda \Delta \underline{\phi}^T \Delta \underline{\phi}$ is:

$$\Delta \underline{\phi} = (C^T C + \lambda I)^{-1} C^T \Delta \underline{X}. \quad (18)$$

After calculating the expression (18), the current guess at the parameters may then be updated using:

$$\underline{\phi} = \underline{\phi}_0 - \Delta \underline{\phi}. \quad (19)$$

Iteratively applying (18) and (19) results in a Levenberg-Marquardt style nonlinear search that should converge if the initial guess is sufficiently close to the true solution (Norton 1986). The free parameter λ is used for numeric conditioning of the matrix inverse in (18). λ may be set to a nonzero value during the iteration, but for convergence of the solution, λ must approach zero. If the matrix inverse $(C^T C)^{-1}$ is undefined near a solution, it means that certain parameters in $\underline{\phi}$ are unidentifiable (see Section 4 for further comments).

Jacobian Calculation

A useful method of calculating the Jacobian matrices in (16) is to consider what instantaneous variations (partial velocities) in the modeled position of the point P are caused by separate small variations in the parameters (Whitney 1972; Bennett and Hollerbach 1988). Because all vectors are functions of the joint angles (at the ith configuration), we will at times explicitly show this by a superscript i. For the purposes of this section the superscript i will be suppressed for notational simplicity (e.g., $z_j^i = z_j$).

For the manipulator kinematic chain model at the ith joint angle configuration, imagine a variation $\Delta \underline{x}^i$ at the point P to be an instantaneous linear velocity vector. The combined variation in all the parameters is presumed to cause this end-point variation. Specifically, a variation of the D-H parameter s_j along the local link z axis z_{j-1} causes a contribution to the end-effector linear velocity of $\Delta s_j z_{j-1}$. Likewise, a parameter variation $\Delta \alpha_j$ about the local link x axis $\boldsymbol{x}_j$ causes a contribution to the end point's linear velocity of $(\Delta \alpha_j \boldsymbol{x}_j) \times \boldsymbol{b}_{j+1}$, where $\boldsymbol{b}_{j+1}$ is a vector from the jth coordinate system to the end point (see Fig. 1). The θ_j and a_j parameters are treated analogously. In total, the end-point translation resulting from all of the parameter variations is given by:

$$\sum_{j=1}^{nf} \mathbf{z}_{j-1} \times \mathbf{b}_j \Delta \theta_j + \mathbf{z}_{j-1} \Delta s_j + \mathbf{x}_j \times \mathbf{b}_{j+1} \Delta \alpha_j + \mathbf{x}_j \Delta a_j. \quad (20)$$

Recalling that $\Delta \underline{x} = [\partial \underline{\hat{x}}^i / \partial \underline{\phi}] \Delta \underline{\phi}$, it is seen that the columns of the Jacobian with respect to each of the parameter variations are given by:

$$\mathrm{col}_j \frac{\partial \underline{\hat{x}}^i}{\partial \underline{a}} = [\mathbf{x}_j], \qquad \mathrm{col}_j \frac{\partial \underline{\hat{x}}^i}{\partial \underline{s}} = [\mathbf{z}_{j-1}] \quad (21)$$

and

$$\mathrm{col}_j \frac{\partial \underline{\hat{x}}^i}{\partial \underline{\boldsymbol{\theta}}} = [\mathbf{z}_{j-1} \times \mathbf{b}_j], \quad \mathrm{col}_j \frac{\partial \underline{\hat{x}}^i}{\partial \underline{\boldsymbol{\alpha}}} = [\mathbf{x}_j \times \mathbf{b}_{j+1}]. \quad (22)$$

The partial derivatives of the camera model may be obtained by similar methods. First concentrate on the length parameters. Notice that the vectors $\mathbf{p}$ and $\mathbf{l}$ in equations (7) and (8) may also be written as:

$$-\mathbf{p} = \sum_{j=0}^{3} \bar{s}_j \bar{\mathbf{z}}_{j-1} + \bar{a}_j \bar{\mathbf{x}}_j, \quad (23)$$

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

$$\mathbf{l} = \sum_{j=2}^{3} \tilde{s}_j \tilde{\mathbf{z}}_{j-1} + \tilde{a}_j \tilde{\mathbf{x}}_j, \tag{24}$$

where $\tilde{\mathbf{z}}$ and $\tilde{\mathbf{x}}$ are local camera z- and x-axes. Thus the camera model of point P (11) can be rewritten to explicitly show the linear dependence of the length parameters:

$$\underline{\tilde{\mathbf{x}}}^i = -\sum_{j=0}^{1} \tilde{s}_j[\Gamma(\tilde{\mathbf{z}}_{j-1})\mathbf{u}_L] + \tilde{a}_j[\Gamma(\tilde{\mathbf{x}}_j)\mathbf{u}_L] + \sum_{j=2}^{3} \tilde{s}_j[\tilde{\mathbf{z}}_{j-1} - \Gamma(\tilde{\mathbf{z}}_{j-1})\mathbf{u}_L] + \tilde{a}_j[\tilde{\mathbf{x}}_j - \Gamma(\tilde{\mathbf{x}}_j)\mathbf{u}_L] \tag{25}$$

where we have taken advantage of the linearity of the operator $\Gamma(\cdot)$ defined in (12). It is thus apparent that the column of the Jacobian with respect to a particular length parameter is given by the term in square brackets that it multiplies in the previous equation. For example, the columns with respect to $\tilde{s}_1$ and $\tilde{s}_2$ are, respectively:

$$\mathrm{col}_j \frac{\partial \underline{\tilde{\mathbf{x}}}^i}{\partial \tilde{s}_1} = [\Gamma(\tilde{\mathbf{z}}_0)\mathbf{u}_L], \quad \mathrm{col}_j \frac{\partial \underline{\tilde{\mathbf{x}}}^i}{\partial \tilde{s}_2} = [\tilde{\mathbf{z}}_1 - \Gamma(\tilde{\mathbf{z}}_1)\mathbf{u}_L]. \tag{26}$$

The other parameters may be treated by direct differentiation of (11). For example, consider the movement of the point P caused by a variation in $\tilde{\alpha}_2$:

$$\Delta \underline{\mathbf{x}}_\alpha = \Delta \tilde{\alpha}_2 \left[\frac{\partial c}{\partial \tilde{\alpha}_2} \mathbf{u}_L + c \tilde{\mathbf{x}}_2 \times \mathbf{b}_3' \right], \tag{27}$$

where $\mathbf{b}_3' = \mathbf{u}_L$. The term in square brackets in (27) is the column of the Jacobian with respect to $\tilde{\alpha}_2$. The evaluation of $\partial c/\partial \tilde{\alpha}_2$ is straightforward but messy. It involves only the partial derivatives of $\mathbf{u}_L$ and $\mathbf{p}$, which are, respectively, $\tilde{\mathbf{x}}_2 \times \mathbf{b}_3'$ and $\tilde{\mathbf{x}}_2 \times \mathbf{p}$.

4. Model Identifiability

If an insufficient number of data points are collected during the calibration or the model used has too many parameters, then various parameters will be unidentifiable. That is, there will be a continuous infinity of possible solutions to the kinematic equations. To quantify this concept, we inspect the first differential of the loop equations (13): $0 = \boldsymbol{C}\Delta\underline{\boldsymbol{\phi}}$. If at or near a correct solution $\underline{\boldsymbol{\phi}}$ there is an infinity of possible solutions for $\underline{\boldsymbol{\phi}}$, then there must be a nonzero solution to the equation $0 = \boldsymbol{C}\Delta\underline{\boldsymbol{\phi}}$ (the null space of $\boldsymbol{C}$). This may only occur if the columns of $\boldsymbol{C}$ (when evaluated at or near the solution $\underline{\boldsymbol{\phi}}$) are linearly dependent. Inspecting the form of $\boldsymbol{C}$ defined in equation (17) shows that the columns of $\boldsymbol{C}$ will be independent and thus the model identifiable if and only if there does not exist some fixed linear relation among the columns of the Jacobians $\boldsymbol{C}^i$. Using (21), (22), (26), and (27), the following condition is obtained:

> **Identifiability condition:** identifiability is guaranteed by checking that there be no constant linear relation (for all $i = 1 \ldots m$ configurations) among the manipulator local link $\mathbf{x}_j^i$-axes and $\mathbf{z}_j^i$-axes, their moment vectors $\mathbf{x}_j^i \times \mathbf{b}_j^i$ and $\mathbf{z}_{j-1}^i \times \mathbf{b}_j^i$, and the camera Jacobian vectors given in the previous section.

This condition has the following physical interpretation: the parameters are identifiable, provided there does not exist a nonzero change in the parameters $\Delta\underline{\boldsymbol{\phi}}$ that induces no change in all of the measured end-point positions. That is, the null space of the $\boldsymbol{C}$ only contains $\Delta\underline{\boldsymbol{\phi}} = 0$.

The identifiability of the parameters of manipulators has previously been discussed in the context of a similar identifiability condition (Bennett and Hollerbach 1989a); thus we will restrict ourselves to camera-related problem situations.

As a first example of unidentifiable parameters, consider the situation when the two camera rotation axes are parallel: $\tilde{\mathbf{z}}_0^i = \tilde{\mathbf{z}}_1^i \equiv \mathbf{z}^i$. One might suspect a problem, because it is well known that the D-H s parameter is poorly defined for serial link kinematic chains with consecutive parallel axes (Hayati 1983). The columns of the Jacobian (26) with respect to $\tilde{s}_1$ and $\tilde{s}_2$ are $\Gamma(\mathbf{z}^i)\mathbf{u}_L^i$ and $\mathbf{z}^i - \Gamma(\mathbf{z}^i)\mathbf{u}_L^i$, respectively. These two vectors are not linearly dependent, so the problem is not as simple as for serial link chains. Notice though that $\tilde{\mathbf{z}}_1^i = \mathbf{z}^i$ is the base coordinate's z-axis of the manipulator, and the Jacobian vector with respect to the manipulator parameter s_0 is $\mathbf{z}$. Thus there exists a linear relation among these three Jacobian vectors:

$$0 = -[\mathbf{z}^i] + [\mathbf{z}^i - \Gamma(\mathbf{z}^i)\mathbf{u}_L^i] + [\Gamma(\mathbf{z}^i)\mathbf{u}_L^i], \tag{28}$$

for all i. Thus by the identifiability condition, the parameters s_0, $\tilde{s}_1$, and $\tilde{s}_0$ are not identifiable alone. This becomes a practical problem, because it also means that $\boldsymbol{C}^T\boldsymbol{C}$ will be singular in the iteration algorithm (18). The solution is to use an alternate coordinate convention for the transformation $\tilde{A}_1$. The Hayati convention (Hayati 1983) developed for manipulators provides such a four-parameter system. This convention cannot be used for all joints, because it, too, has a similar ambiguous parameter when there are two consecutive perpendicular axes.

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

As a second instance of unidentifiable parameters, notice that the closed-loop equation (13) of the calibrated model may be written to show explicitly the linear dependence of the length parameters:

$$\begin{aligned} 0 = &-\sum_{j=0}^{1} \tilde{s}_j[\Gamma(\tilde{\mathbf{z}}_{j-1})\mathbf{u}_L] + \tilde{a}_j[\Gamma(\tilde{\mathbf{x}}_j)\mathbf{u}_L] \\ &+ \sum_{j=2}^{3} \tilde{s}_j[\tilde{\mathbf{z}}_{j-1} - \Gamma(\tilde{\mathbf{z}}_{j-1})\mathbf{u}_L] + \tilde{a}_j[\tilde{\mathbf{x}}_j - \Gamma(\tilde{\mathbf{x}}_j)\mathbf{u}_L] \\ &- \sum_{j=0}^{nf} s_j[\mathbf{z}_{j-1}] + a_j[\mathbf{x}_j] \end{aligned} \tag{29}$$

where, in addition to equations (13) and (26), the fact that the manipulator length parameters enter linearly into the kinematics (Bennett and Hollerbach 1989a) has been used. All of the terms in square brackets are columns of the Jacobian C^i, and thus the closed loop is unidentifiable by the above "identifiability condition." Also, from (29) it is seen that the source of the trouble is that the loop equations may be scaled arbitrarily and still satisfy the joint angle and image plane data. The solution to this problem is to fix one link length arbitrarily. This length determines the units of length. If conventional units, such as a meter, are desired, the calibrated vision system merely has to view the desired meter stick and calculate the correction scale factor to its internal units of length.

As a final example of unidentifiable parameters, consider the case when the data are not "persistently exciting." For instance, if the hand point P is always to move such that it stayed in a plane defined by the right camera axis $\tilde{\mathbf{z}}_0^i$, and this plane happened to be coincident with both focal points, then $(\mathbf{u}_R^i \times \mathbf{u}_L^i) \times \tilde{\mathbf{z}}_0^i = 0$ for all i configurations. Thus $\Gamma(\tilde{\mathbf{z}}_0^i) = 0$, and the column of the Jacobian with respect to $\tilde{s}_1$ is identically zero. The parameter $\tilde{s}_1$ is unidentifiable in this situation. Though this scenario is unlikely, it does point out the importance of the choice of the configuration data used for calibration. Although perhaps only simulation may determine the optimal data set, it is possible to study the sensitivity of particular parameters (Torre et al. 1986).

5. An Example and Simulation

To clarify the general calibration procedure, an example is now presented and solved by simulation. Consider the planar 2-DOF manipulator connected to a head mounted stereo system in Figure 5. In total there are six degrees of freedom in the system:

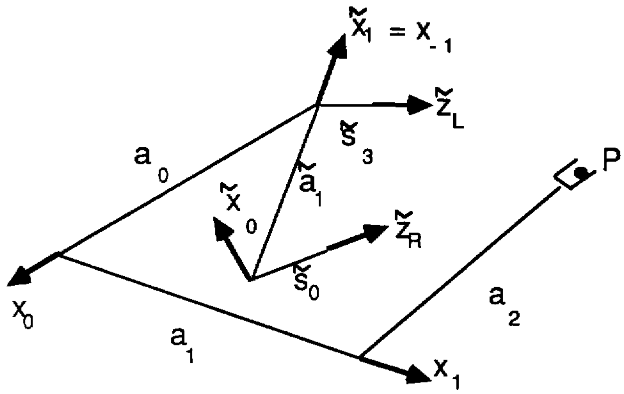

Fig. 5. Example hand-eye system used for simulation purposes.

two manipulator joints, two camera rotations, and two one-dimensional images. The kinematic parameters of interest are placed into the array $\underline{\phi}$ in the following order: the three manipulator link lengths and joint angle offsets, three length parameters providing the displacement between camera coordinates, two camera rotation offsets, and two focal lengths: $\underline{\phi} = (a_0, a_1, a_2, \theta_0, \theta_1^{off}, \theta_2^{off}, \tilde{s}_0, \tilde{a}_1, \tilde{s}_3, \tilde{\theta}_1^{off}, \tilde{\theta}_2^{off}, f_L, f_R)^T$. A simulation of this hand-eye system was performed by using the "actual" parameter values $\underline{\phi} = (1, 1, 1, -0.8, 0, 0, 0.1, 0.36, 0.1, 0, 0, 0.05, 0.05)^T$ (lengths are in meters and angles in radians). Joint angle data were generated by moving the four rotational joints θ_1, θ_2, $\tilde{\theta}_1$, and $\tilde{\theta}_2$ over a trajectory starting at $\theta_1 = 0$, $\theta_2 = 0$, $\tilde{\theta}_1 = 0$, and $\tilde{\theta}_2 = 0$ and covering a joint space volume defined by increasing each joint three times in increments of 0.1 rad (i.e., $3 \times 3 \times 3 \times 3 = 81$) distinct configurations. The resulting joint angles and calculated image pairs were used as input to the algorithm given by equations (18) and (19). In addition, a preliminary guess of $\underline{\phi}_0 = (1.1, 1.1, 1.1, -.9, .1, .1, .11, .4, .11, .1, .1, .06, .06)^T$ was provided, and $\tilde{s}_0$ was fixed at 0.1m. The algorithm was run until the "actual" parameters were recovered to within eight decimal places.

6. Conclusions

A general framework for calibrating a manipulator and stereo system for performing the task of hand-eye coordination has been presented. The emphasis has been on autonomous calibration. The vision system was seen to be represented simply with the D-H convention, thus allowing a unification of manipulator and camera model notation. The vector-based derivation of the columns of the closed-loop Jaco-

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

bian enabled an identifiability condition to be derived. As examples of the application of this condition, the parallel axes problem, the length parameter scale problem, and the data "persistency of excitation" problem were all discussed. Finally, a Newton-like iterative search procedure was prescribed for identifying the kinematic parameters.

Acknowledgments

This article describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the University Research Initiatives under Office of Naval Research contract N00014-86-K-0180 and in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124. Personal support for J. M. Hollerbach was also provided by a National Science Foundation Presidential Young Investigator Award and for D. J. Bennett by a Fairchild fellowship and a scholarship from the Natural Sciences and Engineering Research Council of Canada.

References

Abdel-Aziz, Y. I., and Karara, H. M. 1971. Direct linear transformation into object space coordinates in close-range photogrammetry. Symposium on Close-Range Photogrammetry. Urbana, Ill.: University of Illinois at Urbana-Champaign, pp. 1–18.

An, C. H., Atkeson, C. G., and Hollerbach, J. M. 1988. *Model-based Control of a Robot Manipulator*. Cambridge, Mass.: MIT Press.

Bennett, D. J. 1990. The Control of Human Arm Movement: Models and Mechanical Constraints. Ph.D. thesis. Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology.

Bennett, D. J., and Hollerbach, J. M. 1988 (Austin, Tex., Dec. 7–9). Self-calibration of single-loop, closed kinematic chains formed by dual or redundant manipulators. *Proc. 27th IEEE Conf. Decision and Control*, pp. 627–629.

Bennett, D. J., and Hollerbach, J. M. 1989a (Scottsdale, Ariz., May 14–19). Identifying the kinematics of robots and their tasks. *Proc. IEEE Int. Conf. Robotics and Automation*, pp. 580–586.

Bennett, D. J., and Hollerbach, J. M. 1989b (Columbus, Ohio, June 13–15). Identifying the kinematics of non-redundant serial chain manipulators by a closed-loop approach. *Proc. Fourth International Conference on Advanced Robotics*. In press.

Brooks, R. A., Flynn, A. M., and Marill, T. 1988. Self calibration of motion and stereo for mobile robots. In Bolles, R., and Roth, B. (eds.): *Robotics Research: The Fourth International Symposium*. Cambridge, Mass.: MIT Press, pp. 277–286.

Denavit, J., and Hartenberg, R. S. 1955. A kinematic notation for lower pair mechanisms based on matrices. *J. Appl. Mechanics* 22:215–221.

Duda, R. O., and Hart, P. E. 1973. *Pattern Recognition and Scene Analysis*. New York: Wiley.

Ganapathy, S. 1984 (Altanta). Decomposition of transformation matrices for robot vision. *Proc. IEEE Int. Conf. Robotics and Automation*, pp. 130–139.

Gremban, K. D., Thorpe, C. E., and Kanade, T. 1988 (Philadelphia, April 24–29). Geometric camera calibration using systems of linear equations. *Proc. IEEE Int. Conf. Robotics and Automation*, pp. 562–567.

Haralick, R. M. 1980. Using perspective transformations in scene analysis. *Comp. Graphics Image Process.* 13:191–221.

Hayati, S. A. 1983 (San Antonio, Dec. 14–16). Robot arm geometric link parameter estimation. *Proc. 22nd IEEE Conf. Decision and Control*, pp. 1477–1483.

Hollerbach, J. M. 1989. A survey of kinematic calibration. In Khatib, O., Craig, J. J., and Lozano-Pérez, T. (eds.): *Robotics Review 1988*. Cambridge, Mass.: MIT Press.

Horn, B. K. P. 1986. *Robot Vision*. Cambridge, Mass.: MIT Press.

Isaguirre, A., Pu, P., and Summers, J. 1985 (St. Louis, March 25–28). A new development in camera calibration: Calibration of a pair of mobile cameras. *IEEE Int. Conf. Robotics and Automation*, pp. 74–79.

Kenefick, J. F., Gyer, M. S., and Harp, B. F. 1972. Analytic self-calibration. *Photogrammetr. Eng.* 38:1117–1126.

Martins, H. A., Birk, J. R., and Kelley, R. B. 1981. Camera models based on data from two calibrated planes. *Comp. Graphics Image Process.* 17:173–180.

Moffitt, F., and Mikhail, E. 1980. *Photogrammetry*. New York: Harper and Row.

Niemeyer, G., and Slotine, J-J. E. 1988 (Austin, Tex., Dec. 7–9). Performance in adaptive manipulator control. *Proc. 27th IEEE Conf. Decision and Control*, pp. 1585–1591.

Norton, J. P. 1986. *An Introduction to Identification*. Orlando, Fla.: Academic Press.

Puskorius, G. V., and Feldkamp, L. A. 1987 (Raleigh, N.C., March 31–April 3). Global calibration of a robot/vision system. *Proc. IEEE Int. Conf. Robotics and Automation*, pp. 190–195.

Sobel, I. 1974. On calibrating computer controlled cameras for perceiving 3-D scenes. *Artif. Intell.* 5:185–198.

Tsai, R. Y. 1986 (Miami). An efficient and accurate camera calibration technique for 3-D machine vision. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 364–374.

Tsai, R. Y. 1989. Synopsis of recent progress on camera calibration for 3D machine vision. In Khatib, O., Craig, J. J., and Lozano-Pérez, T. (eds.): *Robotics Review 1988*. Cambridge, Mass.: MIT Press.

Torre, V., Verri, A., and Fiumicelli, A. 1986. Stereo accu-

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

racy for robotics. In Faugeras, O., and Giralt, G. (eds.): *Robotics Research: The Third International Symposium.* Cambridge, Mass.: MIT Press, pp. 5–10.
Whitney, D. E. 1972. The mathematics of coordinated control of prosthetic arms and manipulators. *ASME J. Dyn. Sys. Meas. Control* 303–309.
Yakimovsky, Y., and Cunningham, R. 1978. A system for extracting three dimensional measurements from a stereo pair of TV cameras. *Comp. Graphics Image Process.* 7:195–210.
Ziemann, H. 1985 (Washington, D.C., March 10–15). A mathematical model for camera calibration. *Technical Papers, 51st Annual American Society of Photogrammetry Meeting*, pp. 725–735.

Downloaded from http://ijr.sagepub.com at UNIV OF UTAH on November 19, 2007
© 1991 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.