Propagation and reduction of error in three-dimensional structure determined from biplane views of unknown orientation

Laura E. Fencil
Department of Radiology, The University of Michigan, Ann Arbor, Michigan 48109

Charles E. Metz[a)]
Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, Chicago, Illinois 60637

(Received 19 June 1989; accepted for publication 31 July 1990)

We are developing a technique for determination of the three-dimensional (3-D) structure of vascular objects from two radiographic projection images acquired at arbitrary and unknown relative orientations. No separate calibration steps are required with this method, which exploits an inherent redundancy of biplane imaging to extract the imaging geometry as well as the 3-D locations of eight or more object points. The theoretical basis of this technique has been described previously. In this paper, we review the method from the perspective of linear algebra and describe an improvement, not heretofore reported, that reduces the method's sensitivity to experimental error. We then examine the feasibility and inherent accuracy of this approach by computer simulation of biplane imaging experiments. The precision with which 3-D object structure may be retrieved, together with the dependence of precision on the actual imaging geometry and errors in various measured quantities, is studied in detail. Our simulation studies show that the method is not only feasible but potentially accurate, typically determining object-point configurations with root-mean-square (RMS) error on the order of 1 to 2 mm. The method is also quite fast, requiring approximately one second of CPU time on a VAX 11/750 computer (0.6 MIPS).

Key words: biplane radiography, angiography, three-dimensional vascular structure, x-ray and γ-ray instruments and techniques, diagnostic methods and instrumentation

I. INTRODUCTION

Techniques that allow determination of the three-dimensional (3-D) structure of cerebral or coronary vasculature have been investigated for several years. Automated image analysis has been applied to stereoscopic digital angiographic images to produce 3-D models of sections of cerebral vessels and to make quantitative measurements of the vessels and flow rate, for example.[1-4] Other workers have reported techniques that utilize biplane digital angiographic images obtained in exactly orthogonal directions.[5-8] MacKay and colleagues[9,10] showed the feasibility of determining 3-D object structure from biplane images obtained at arbitrary orientations when images of a calibration object are obtained under identical conditions. But despite the demonstrated feasibility of these techniques, they have not found widespread clinical use, perhaps due to the constraints imposed by stereoscopic or orthogonal imaging geometry in the former methods and the awkwardness of the additional calibration-imaging step required by the latter.

Several investigators working outside the field of medical imaging[11-13] have reported the basis of a technique that allows 3-D object structure to be determined from biplane images obtained at arbitrary relative orientations *without* the use of a calibration object. In a previous publication,[14] we described the theoretical aspects of the approach as it applies to biplane radiography and noted some potential problems in applying it to actual biplane angiographic images.

This method of determining 3-D object structure from two arbitrarily oriented biplane images without the need for calibration images is based on what we have referred to as the *inherent redundancy of biplane imaging*.[14] For every object point of unknown location that appears in both images, there are three unknown quantities—the spatial coordinates of the point (x_i,y_i,z_i)—but four known quantities—the coordinates of that object point in the first and second images, (u_i,v_i) and (u_i',v_i'). Thus, from the image coordinates of each object point of unknown 3-D location, it should be possible to glean one piece of information concerning the configuration, or relative geometry, of the biplane imaging system. In general, the relative geometry of two biplane views can be described by a 3×3 rigid rotation matrix $[R]$, which contains nine elements, and a three-dimensional translation vector $\mathbf{t}$, which contains three. Of the nine elements of the $[R]$, only five are truly independent, leaving a total of eight quantities that must be determined in order to describe the geometry of the biplane imaging system. Therefore, the biplane image coordinates of eight object points of unknown 3-D location are sufficient to determine the relative geometry of the biplane imaging system if some basic information concerning the individual views is known.[14] Once the geometry of the imaging system has been found, determination of the 3-D structure of the object points appearing in both views is straightforward.

 0094-2405/90/060951-11$01.20

The work reported here examines the feasibility of this technique for determining 3-D object structure from biplane images without knowledge of the relative geometry of the two views. Computer simulation studies were performed to study the dependence of the method's precision on the actual biplane geometry, and to examine the sensitivity of the method to errors in information concerning the individual views and in the measured image coordinates of the object points. We implemented the basic approach described previously[14] with an important modification that reduces the adverse effects of experimental error.

II. THEORY

The theoretical basis of our approach is described in detail elsewhere.[14] Here we review briefly the fundamental mathematics of the method from the perspective of linear algebra and describe an improvement, not heretofore reported, that reduces the method's sensitivity to error.

We begin by considering a single view in biplane imaging and defining a "uv" image-coordinate system for it. Unit vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ in the image plane are chosen arbitrarily such that their vector cross product $\hat{\mathbf{w}} \equiv \hat{\mathbf{u}} \times \hat{\mathbf{v}}$ points away from the x-ray focal spot. The origin of the resulting uvw coordinate system is located in the image plane at the point where a perpendicular from the focal spot intersects the image plane, as shown in Fig. 1. A right-handed object coordinate system xyz for the view is then defined such that its origin is located at the focal spot and its axes, specified by unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$, are aligned with the unit vectors $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$ and $\hat{\mathbf{w}}$, respectively. Coordinate systems $u'v'w'$ and $x'y'z'$ for the second imaging view may be defined analogously, as shown in Fig. 1. With coordinate systems defined in this way, we denote by (u_i,v_i) and (u_i',v_i') the image coordinates in the two views of an ith object point located at an unknown position (x_i,y_i,z_i) in the xyz system and at (x_i',y_i',z_i') in the $x'y'z'$ system, respectively.

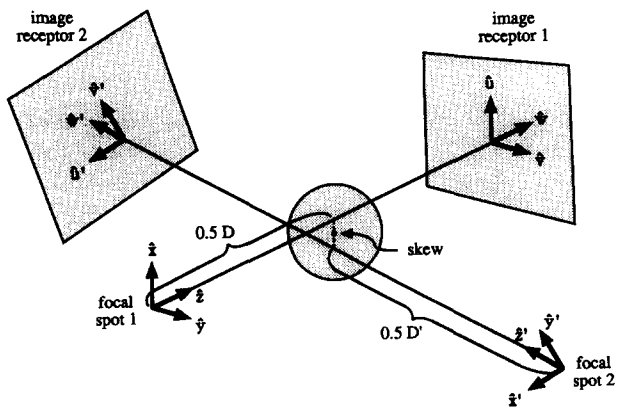


FIG. 1. An example of biplane imaging geometry used in the computer simulation experiments. Spatial coordinate systems xyz and $x'y'z'$ and image coordinate systems uvw and $u'v'w'$ are indicated. In the example shown here, the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes approach most closely at ($z = 0.5$ D, $z' = 0.5$ D'), and skew is present. Object points generated in the simulations were required to fall both within the common volume of the two views and within a 10-cm sphere centered at the intersection (or point of closest approach) of the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes.

Since both the xyz and $x'y'z'$ coordinate systems are right-handed, the transformation between them is defined by a translation and rigid rotation, and hence may be expressed as

$$\begin{bmatrix} x_i' \\ y_i' \\ z_i' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \left\{ \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} - \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \right\}, \tag{1}$$

where the r_{jk} are elements of a 3×3 unitary matrix $[R]$ with determinant equal to $+1$, and where t_x, t_y and t_z are the elements of a vector $\mathbf{t}$ that describes, in the xyz system, the translation from the first focal spot position to the second. Because the three linear equations represented by Eq. (1) specify the relationship between the xyz and the $x'y'z'$ coordinate systems, they link the two image coordinate systems if the following basic information concerning the individual projection views is known: (i) the perpendicular distances (D and D' in Fig. 1) between the focal spots and their image planes; and (ii) the point at which the perpendicular from each focal spot intersects its image plane, which serves as the origin of corresponding image coordinate system.

Inspection of Fig. 1 shows that the image coordinates of an ith object point are given by $(u_i,v_i) = (Dx_i/z_i, Dy_i/z_i)$ in the first view and by $(u_i',v_i') = (D'x_i'/z_i', D'y_i'/z_i')$ in the second view. For convenience in the development that follows, we define dimensionless scaled coordinates $(\xi_i,\eta_i) \equiv (x_i/z_i, y_i/z_i) = (u_i/\mathrm{D}, v_i/\mathrm{D})$ in the first image and $(\xi_i',\eta_i') \equiv (x_i'/z_i', y_i'/z_i') = (u_i'/\mathrm{D}', v_i'/\mathrm{D}')$ in the second image. If the distances D and D' are known, then these scaled coordinates can be determined directly from the measured image coordinates in the two views. For each object point identified in both images, this allows us to construct four equations that relate the scaled image coordinates to the (generally unknown) object point coordinates in the xyz and $x'y'z'$ coordinate systems. Then by use of Eq. (1), which relates the $x'y'z'$ and xyz systems, we obtain four equations involving the scaled image coordinates (ξ_i,η_i) and (ξ_i',η_i'), the object point coordinates (x_i,y_i,z_i), $[R]$, and $\mathbf{t}$.

Solving these four equations for x_i, y_i and z_i yields a single equation involving $[R]$, $\mathbf{t}$, and the scaled image coordinates that may be expressed concisely as

$$[\xi_i' \quad \eta_i' \quad 1][Q] \begin{bmatrix} \xi_i \\ \eta_i \\ 1 \end{bmatrix} = 0, \tag{2}$$

in which $[Q]$ is the 3×3 matrix

$$[Q] \equiv \begin{bmatrix} (r_{13}t_y - r_{12}t_z) & (r_{11}t_z - r_{13}t_x) & (r_{12}t_x - r_{11}t_y) \\ (r_{23}t_y - r_{22}t_z) & (r_{21}t_z - r_{23}t_x) & (r_{22}t_x - r_{21}t_y) \\ (r_{33}t_y - r_{32}t_z) & (r_{31}t_z - r_{33}t_x) & (r_{32}t_x - r_{31}t_y) \end{bmatrix}. \tag{3}$$

Thus, a single expression with the form of Eq. (2) is obtained from the biplane image coordinates of each object point of unknown position that can be located in both views. With N object points located in both biplane images, the resulting system of equations can be written as

$$[A]_{N\times 9}\begin{bmatrix} q_{1,1} \\ q_{1,2} \\ \vdots \\ q_{3,3} \end{bmatrix}_{9\times 1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{N\times 1}, \tag{4}$$

in which each row of $[A]$ represents the coefficients of the elements of $[Q]$ in Eq. (2) and the 3×3 matrix $[Q]$ is expressed now as a 9×1 column vector. Since Eq. (4) is homogeneous, its solution can be determined only to within an unknown multiplicative constant. Hence, for convenience, we are free to scale any solution $[Q^*]$ so that the sum of its squared elements is equal to unity, which is equivalent to normalizing the 9×1 column-vector version of $[Q^*]$ to be a unit vector. When error is absent, the elements of this unit vector in nine-dimensional space are related to those of the true $[Q]$ by an unknown multiplicative constant.

The exact solution of Eq. (4), when one exists—and the least-squares solution when no exact solution exists—is given by the unit eigenvector that corresponds to the smallest of the nine eigenvalues of $[A]^T[A]$. Hereafter we denote the ith eigenvalue of this matrix, in order of decreasing magnitude, by λ_i. Recall that the eigenvectors of *any* matrix product of the form $[A]^T[A]$ represent the principal axes of the rows of $[A]$, and that each eigenvalue represents the sum of the squared components of those row vectors in the direction of the corresponding principal axis.

With $N = 8$ object points, or with $N > 8$ object points if experimental and computational error is absent [i.e., if the $N > 8$ equations of Eq. (4) are consistent], the rows of $[A]$ will lie in an eight-dimensional subspace or hyperplane. In this situation, there necessarily exists a solution vector $[Q^*]$ such that its scalar product with each of the rows of $[A]$ is equal to zero. Thus, the ninth, or smallest, eigenvalue of $[A]^T[A]$ must equal zero, and the corresponding eigenvector provides an exact solution to Eq. (4). Aside from its sign and magnitude, this exact solution is unique (unless the points lie in certain "degenerate" 3-D configurations that make some rows of $[A]$ linearly dependent[12]). However, if experimental and/or computational error is *not* negligible, then with more than eight object points the rows of $[A]$ generally will deviate slightly from a common 8-dimensional hyperplane. In this situation, the unit eigenvector corresponding to the smallest eigenvalue of $[A]^T[A]$—which we take as our solution $[Q^*]$—corresponds to the direction in nine-dimensional space that is perpendicular to the hyperplane which best fits the rows of $[A]$ in a least-squares sense. The amount by which the rows of $[A]$ deviate from this best-fit hyperplane is expressed by the mean squared projection, λ_9/N, which represents the variance of the row vectors in the direction perpendicular to the best-fit hyperplane. Thus, in situations involving more than eight object points, one may expect that the magnitude of λ_9 should reflect the magnitude of experimental and computational error. In other words, as λ_9/N (or λ_9, for a fixed number of points) increases, the average error of the solution may be expected to increase as well.

The stability of the perpendicular to the best-fit hyperplane in nine-dimensional solution space depends upon the stability of the best-fit hyperplane, even if all N rows of $[A]$ lie within it. When the *true* row vectors of $[A]$ are almost parallel in their common hyperplane, the orientation of that hyperplane is more sensitive to error in the vectors' elements (i.e., in the image coordinates) than when the row vectors fan out broadly. The degree to which the row vectors diverge within the best-fit hyperplane is reflected by the ratio λ_1/λ_8, which we refer to henceforth as the "condition number" of $[A]^T[A]$. Larger condition numbers indicate rows of $[A]$ that are more nearly collinear in their best-fit hyperplane, and so should predict situations in which the solution of Eq. (4) is more sensitive to error in measured image coordinates.

We described elsewhere[14] a technique by which the solution $[Q^*]$ may be improved, in the presence of experimental error, by taking note of the fact that the three rows of the *true* 3×3 matrix $[Q]$, which can be interpreted as 3-D vectors $\mathbf{q}_k$, must be coplanar in 3-D space. The improved solution $[Q^{**}]$ that is obtained by forcing the rows of $[Q^*]$ to occupy a "best-fit" common plane can be scaled so it is appropriate to a translation vector of unit magnitude, thereby producing a scaled solution $[Q^{***}]$.[14] We also showed that an estimate of the direction of $\mathbf{t}$, which is a unit vector and is designated $\hat{\mathbf{t}}^*$, may be obtained from the normal to this common plane.[14] However, we had been unable to ensure that the scaled solution $[Q^{***}]$ would result in an estimate of $[R]$ with the properties of a rigid rotation matrix (unitary with determinant equal to $+1$). Here we report a simple procedure which provides a solution that ensures such an estimate of $[R]$. Although the improved solution $[Q^{****}]$ can be calculated from the solution $[Q^{***}]$ discussed previously, it can also be defined directly in terms of the initial (i.e., unscaled and uncorrected) solution $[Q^*]$ as

$$[Q^{****}] \equiv [Q^*][E][D][E]^T, \tag{5}$$

where $[D]$ is the diagonal matrix

$$[D] \equiv \begin{bmatrix} \frac{1}{\sqrt{\varphi_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\varphi_2}} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{6}$$

in which φ_1 and φ_2 are the largest and second-largest eigenvalues of $[Q^*]^T[Q^*]$, respectively, and where $[E]$ is a 3×3 unitary matrix with columns $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, and $\hat{\mathbf{e}}_3$ that are the unit eigenvectors corresponding to φ_1, φ_2 and φ_3, respectively. In effect, the process summarized by Eq. (5) forces the eigenvalues of $[Q^{****}]^T[Q^{****}]$ to be $\{+1, +1, 0\}$, which is a necessary and sufficient condition for the rows of $[Q^{****}]$ to represent projections of the rows of a rigid rotation matrix onto a plane. With this procedure, the estimate $\hat{\mathbf{t}}^*$ provided by $\hat{\mathbf{e}}_3$ is orthogonal to the rows of $[Q^{****}]$, so it can be used with the rows of $[Q^{****}]$ to calculate the rows of an estimate of $[R]$ according to[14]

$$\mathbf{r}_1 = (\mathbf{q}_1^{****} \times \hat{\mathbf{t}}^*) + (\mathbf{q}_2^{****} \times \mathbf{q}_3^{****}), \text{ etc.} \tag{7}$$

One can show that of all estimates of $[Q]$ which in this way produce an $[R]$ with the properties of a rigid rotation matrix, $[Q^{****}]$ is the particular estimate which minimizes $\Sigma_{k=1}^{3} |q_k^{****} - q_k^*|^2$. Thus, $[Q^{****}]$ constitutes the valid

solution for $[Q]$ that is closest to the unscaled and uncorrected solution $[Q^*]$ in a least-squares sense. (After completing the work reported here, we learned that Weng, Huang, and Ahuja[15] recently proposed an alternative procedure, based on an analysis of quaternions, that satisfies this condition in the essentially equivalent task of inferring object motion and structure from two fixed-perspective views taken at different times.)

There are in fact four possible combinations of $[R]$ and $\hat{\mathbf{t}}$ that are consistent with $[Q^{****}]$, since the signs of both $[Q^{****}]$ and $\hat{\mathbf{t}}^*$ are ambiguous, but only one combination yields positive values for all z_i and z_i'.[11] Hence, the correct combination of $[R]$ and $\hat{\mathbf{t}}^*$ can be found by calculating the 3D coordinates of the object points with each of the four possible combinations of $[R]$ and $\hat{\mathbf{t}}^*$, and then selecting the combination that yields positive values for all z_i and z_i'.[11,14] Because the *magnitude* of the original translation vector $\mathbf{t}$ cannot be determined with this approach, the calculated 3-D coordinates of the object points are expressed in units of the length of the actual translation vector. The absolute scale of the object can be recovered if the absolute distance between the two focal spot positions or between two or more object points is known, however.[14]

III. SIMULATION STUDY DESIGN

Our computer simulations of biplane imaging were conducted with a VAX 11/750 computer (Digital Equipment Corporation, Maynard, New Jersey), VAX Fortran 77, and double precision floating-point arithmetic. The simulation experiments consisted of the following steps: (i) specification of a biplane imaging geometry; (ii) random placement of a preselected number of infinitesimal "object points" within the common volume of the biplane imaging system (defined as the 3-D region of object space seen by the image receptors of both views); (iii) use of the resulting image coordinates of the object points, along with the specified distances D and D', to determine the imaging geometry; (iv) calculation of the object-point positions from the imaging geometry and the image coordinates; and finally, (v) comparison of the calculated 3-D object-point configuration with the true object-point configuration. Each of these steps is described in detail next.

Several parameters must be specified to define completely the biplane imaging geometry of a particular simulation experiment. The size of the image receptor and the number of pixels in two orthogonal directions determine the effective pixel size of the system. In all of the simulations reported here, we assumed a square 17.8×17.8 cm image receptor, a 512×512 image matrix, and thus a pixel size of 0.035 cm. This combination of parameters is similar to that provided by the 17-cm (7-in.) image intensifier (II) mode of many digital angiographic systems. The perpendicular distances between the two focal spots and their respective image planes, D and D', must be specified also. We set both D and D' equal to 100 cm in our simulation experiments, except where noted otherwise below.

The angular orientation of the second biplane view with respect to the first can be defined by three angles Θ_x, Θ_y, and Θ_z that represent successive rotations about the $\hat{\mathbf{x}}$, rotated $\hat{\mathbf{y}}$, and twice-rotated $\hat{\mathbf{z}}$ axes such that the xyz coordinate system becomes aligned with the $x'y'z'$ system. The point of intersection (or closest approach) of the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes, shown in Fig. 1, is defined by the fractions of the lengths of D and D' at which the intersection (or closest approach) occurs. The amount of skew, if any, between the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes (see Fig. 1) must be specified also. Error as a function of these parameters was investigated in our study.

Following the specification of a particular biplane imaging geometry, a predetermined number of infinitesimal "object points" ($N \geqslant 8$) are scattered within the common volume of the biplane imaging system by use of a uniform-deviate random-number generator. In our work, a constraint is placed on the object-point locations such that they are required to fall both within the common volume of the imaging system and within a sphere of diameter 10 cm that is centered on the point at which the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes intersect. If skew is included in the biplane geometry, then the sphere is centered midway between the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes at their point of closest approach, as shown in Fig. 1. Thus, the object points lie within a volume similar to that of the adult human heart.

After the N object points are generated, the exact image coordinates of each object point in each view, (u_i,v_i) and (u_i',v_i'), are calculated. Unless noted otherwise below, the assumed pixel size is then used to calculate the image coordinates in each view in terms of pixels, after rounding to the nearest pixel. The scaled image coordinates are calculated from these discrete coordinates, and then the elements of the matrix $[A]$ [in Eq. (4)] are determined by combining the scaled image coordinates of the two views appropriately. Subsequently, the eigenvalues and eigenvectors of the 9×9 matrix $[A]^T[A]$ are found, and the eigenvector corresponding to the smallest eigenvalue is taken as the initial solution $[Q^*]$ from which the unit translation vector $\hat{\mathbf{t}}^*$ is determined. The final solution $[Q^{****}]$ is calculated from Eq. (5), and then Eq. (7) is used to extract the rotation matrix $[R]$ from $[Q^{****}]$ and $\hat{\mathbf{t}}^*$. Finally, the 3D coordinates of each object point in the xyz system, (x_i,y_i,z_i), are computed by least squares and converted, for comparison with the true coordinates, to the scale of the original object by multiplying each coordinate by the distance between the two focal-spot positions of the true biplane geometry.

Experience has shown that the resulting estimates of the object-point coordinates, when expressed in the xyz system, can differ by a translation, rotation, and/or scale factor from the true xyz coordinates defined in the simulation, even when the *relative* positions of the object points are recovered accurately. Translation, rotation, and scale-factor errors in estimates of object-point configurations will be of little or no importance in many potential applications of the technique, however, because the 3-D positions of the object points relative to each other (rather than relative to one of the focal spots and its image plane) will be of primary interest, and because the absolute scale of the object is inherently ambiguous when the true distance between the focal-spot positions is unknown, as noted above. Therefore, in order to measure meaningfully the precision with which the object-point configurations are recovered, it is necessary first to translate, rotate, and scale the computed configuration so as to mini-

mize the sum of squared distances between the original and estimated 3-D coordinates of the set of object points. Only then is the accuracy of the estimated configuration calculated, in terms of the minimized root-mean-square (RMS) distance between the true and estimated object-point positions. The condition number of the matrix $[A]^T[A]$ is also computed for each configuration of object points.

In the simulated experiments we varied systematically and individually each parameter of the biplane imaging geometry to investigate the effects of the system geometry on the precision with which the 3-D object-point positions could be recovered. We investigated also the effects of error in measuring the distances D and D′ and the origins of the two image-coordinate systems. Finally, we examined the effects of the number of object points (N) on the precision of the method and explored the effects of errors in the measurement of the image coordinates.

Unless otherwise noted below, we studied fifty independently generated object-point configurations for each combination of imaging geometry, measurement error and N, and averaged RMS distance error and condition number over the fifty configurations. A variety of different configurations of object points was studied for each combination of imaging parameters to reduce the impact of any particularly "well-conditioned" or "ill-conditioned" configuration on the results.

After image coordinates had been determined, calculation of the 3-D object-point coordinates with the VAX 11/750 (0.6 MIPS) usually required less than one second.

IV. RESULTS

A. Effects of imaging geometry and number of object points

The first goal of the simulation study was to examine the inherent ability of the technique to recover 3-D information under "perfect" conditions. To this end, the image coordinates of each object point were computed exactly, rather than simply to the nearest pixel, and no measurement error was introduced into either the origins of the image-coordinate systems or D and D′. Arbitrarily, a single biplane geometry was selected with the following characteristics: both D and D′ were equal to 100 cm; Θ_x, Θ_y and Θ_z were equal to 70°, −70° and 70°, respectively; and the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes intersected at $z = 0.5$ D and $z' = 0.5$ D′, with no skew. The number of object points (N) was varied systematically from eight to thirty, with fifty configurations generated randomly and then analyzed for each value of N. The results are shown in Fig. 2(a), where each datum corresponds to the average RMS distance error obtained from the fifty configurations. In this and all subsequent simulations, the RMS distance error was obtained by comparing the estimated 3-D object-point coordinates with the true 3-D coordinates after the calculated coordinates had first been multiplied by the true length of the translation vector and then translated, rotated, and scaled so as to minimize the RMS distance error between the original and calculated coordinates.

It is clear from Fig. 2(a) that the inherent accuracy of the method under "perfect" conditions was outstanding, with

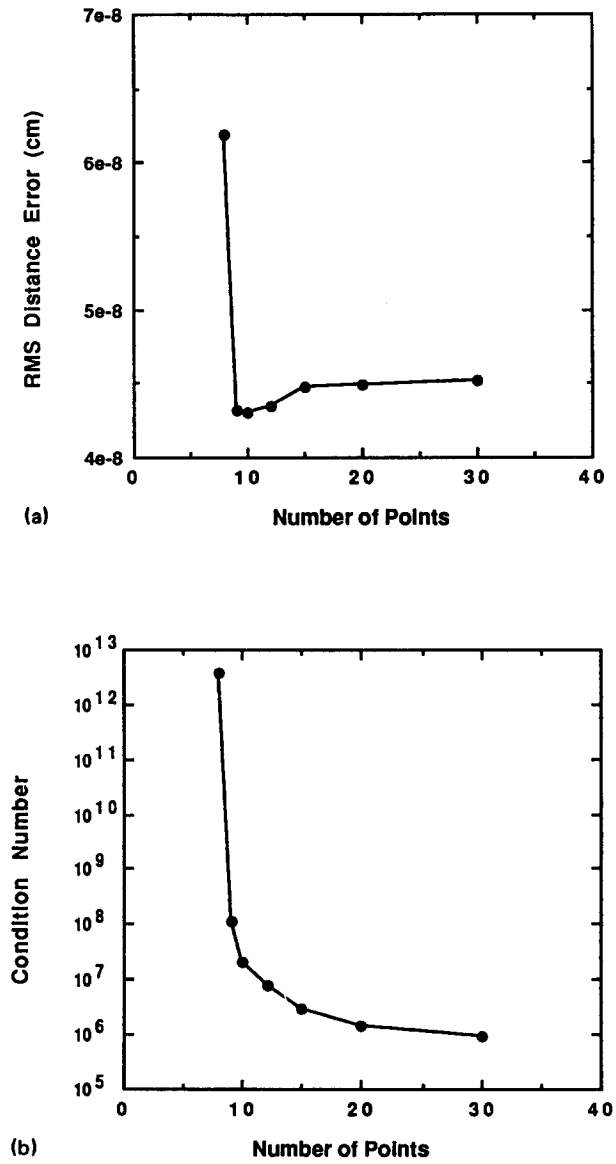


FIG. 2. Effect of the number of object points (N) on RMS 3D reconstruction error and condition number: ideal case. (a) Dependence of RMS error on N. (b) Dependence of the condition number of $[A]^T[A]$ on N. The conditioning of $[A]^T[A]$ improves as the value of N increases.

average RMS 3-D distance error on the order of 10^{-8} cm. (It should be noted that the amount of translation, rotation and scaling of the calculated coordinates that was required in this set of experiments was negligible.) Average error fell approximately 30%, from 6.2×10^{-8} to 4.3×10^{-8}, when the number of object points (N) increased from eight to nine, and then rose slightly as N increased further from nine to thirty. The latter small increase in error may be attributed to an increase in computational roundoff error that occurs as N becomes large. The rather dramatic decrease in error as N went from eight to nine is explained in part by Fig. 2(b), which illustrates the effect of N on the condition number of $[A]^T[A]$. Two things should be noticed from this figure: first, the condition numbers were quite large overall; and second, the average condition number decreased by approximately five orders of magnitude as N increased from eight to

nine, with an additional drop of only two orders of magnitude as N increased further from nine to thirty. Thus, the inherent conditioning of the 3-D reconstruction problem improved dramatically as N increased from eight to nine, leading to the observed decrease in average error. The additional improvement in condition number as N increased from nine to thirty was not sufficient to offset the increasingly important effects of computational roundoff error, however.

To assess the reliability of the technique under somewhat more realistic conditions, we simulated a set of experiments similar to those described above, but with the image coordinates of each object point known only to the nearest pixel. A pixel size of 0.035 cm was used in these and all subsequent simulations. The imaging geometry was the same as that described above; no errors were introduced into D and D′ or the origins of the image coordinate systems; and fifty configurations of object points were examined for each number of object points (N). The dependence of image reconstruction variance (defined as the square of the RMS distance error for each object-point configuration) and the condition number of $[A]^T[A]$ on N is shown in Fig. 3. RMS reconstruction error (the square root of variance) dropped by more than an order of magnitude, from approximately 0.5 to 0.03 cm, as N increased from eight to thirty. It is evident that condition number was a fairly good predictor of 3D reconstruction error in these experiments.

The effect of imaging system geometry on the overall precision of the calculated object coordinates was examined in several ways. The importance of the relative rotation between the two imaging views was assessed by varying the values of Θ_x, Θ_y and Θ_z jointly from 1° to 179°. In this set of experiments D and D′ were set equal to 100 cm; the $\hat{z}$ and $\hat{z}'$ axes intersected at 0.5D and 0.5D′ with no skew in the system, N was equal to ten, and the image coordinates were known to the nearest pixel. The results are shown in Fig. 4. It is clear that RMS error increased substantially as the values of Θ_x, Θ_y and Θ_z approached the extremes of 0° and 180°. This was to be expected, since the two images become more similar as the angles Θ_x, Θ_y and Θ_z approach these extremes. Notably, RMS error is almost independent of angle for rotations between about 40° and 140°.

Figure 5 shows the dependence of RMS error on the location of the intersection of the $\hat{z}$ and $\hat{z}'$ axes when skew was

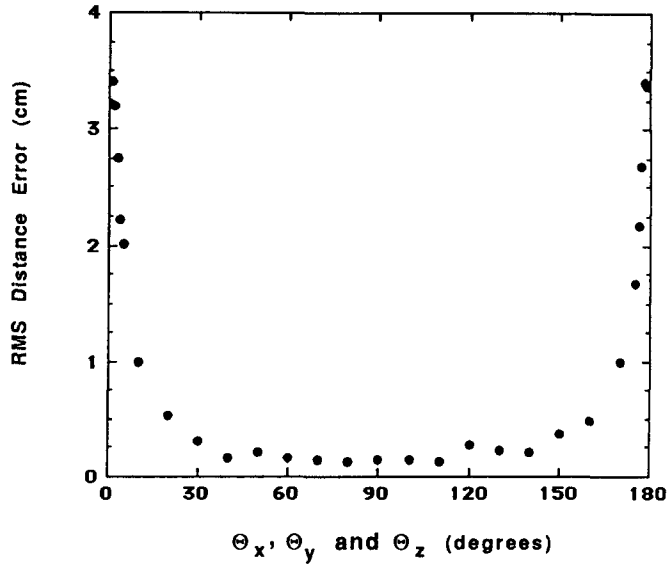


FIG. 4. Effect of the axis-rotation angles Θ_x, Θ_y, and Θ_z on RMS error in the calculated configurations of ten object points. The abscissa indicates rotations in all three angles. Reconstruction error is essentially constant for angles between about 40° and 140°.

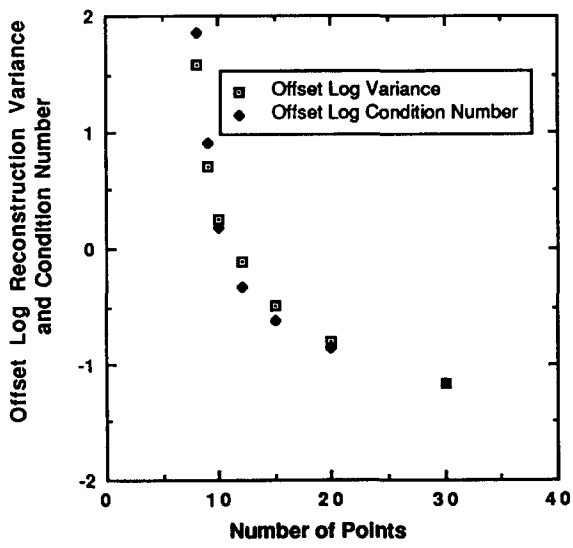


FIG. 3. Dependence of the reconstruction variance, defined as the square of the RMS error, and the condition number of $[A]^T[A]$ on the number of object points (N). The logarithms of the condition number and variance have been offset by their respective mean values to facilitate comparison. Image coordinates were known only to nearest 0.035 cm pixel.

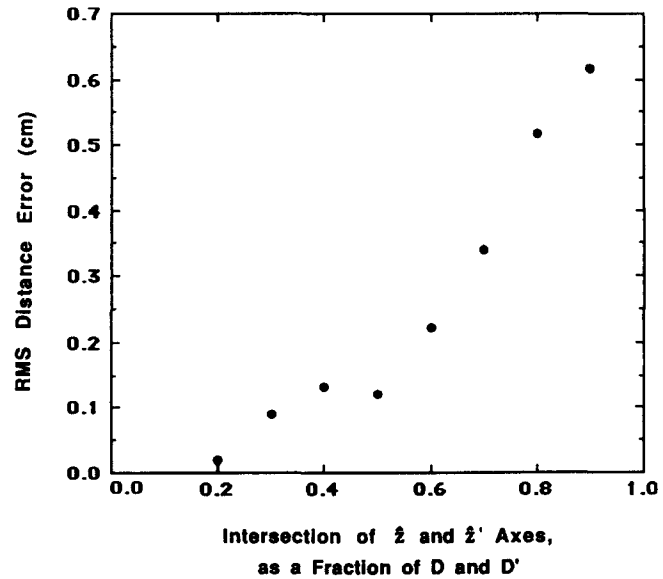


FIG. 5. Effect of the position of the intersection of the $\hat{z}$ and $\hat{z}'$ axes on RMS reconstruction error for $N = 10$; D and D′ = 100 cm; Θ_x, Θ_y, and $\Theta_z = 70°$, $-70°$ and 70°, respectively; and no skew. Average RMS error increases as the intersection moves away from the focal spots.

absent. The point of intersection of the two axes ranged from (0.2D, 0.2D′) to (0.9D, 0.9D′). The general increase in RMS error seen as the point of intersection moved away from the focal spots can be ascribed to two effects. First, the overall "size" of the collection of object points became larger as the point of intersection moved away from the focal spots, because the common volume of the two views became larger. This continued until the common volume exceeded that of the 10-cm-diam sphere into which the points were required also to fit. Second, as the object points moved closer to the image planes, each nearest-pixel quantization error in the image coordinates corresponded to a greater 3-D distance error in the associated object-point position.

We examined the dependence of RMS error on the values of D and D′ by varying these two distances jointly from 70 to 130 cm. In these studies the values of Θ_x, Θ_y and Θ_z were 70°, − 70° and 70°, respectively; the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes intersected at (0.5 D, 0.5 D′) with no skew, and N was equal to ten. Average error increased as both D and D′ became larger, from approximately 0.1 cm at distances of 70 to 0.2 cm at distances of 130 cm. This effect was due to the fact that the solid angle subtended by the collection of object points became smaller as D and D′ lengthened. This caused the rows of $[A]$ to become more nearly equal; thus, the condition number of $[A]^T[A]$ became larger, producing an increasingly ill-conditioned 3-D reconstruction task. A similar phenomenon occurred when the skew between the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes was increased from the usual 0 cm to a value of 6 cm, thus decreasing the common volume of the system. The distances D and D′ were equal to 100 cm in these experiments, and the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes intersected at (0.5 D, 0.5 D′). In this situation, RMS error increased from approximately 0.12 to 0.53 cm. Under these conditions 6 cm of skew substantially reduced the common volume within which the object points could be scattered, so the solid angle subtended by the collection of object points was reduced as well.

B. Effects of error in measured quantities

To implement our technique with actual imaging systems, it is necessary to determine the distances D and D′, the locations of the origins of the image coordinate systems in both image planes, and the coordinates of the N object points in both images. We investigated the effects of error in each of these measured quantities on the accuracy of the method.

The values of D and D′ and the pixel size are used only to form scaled image coordinates, (ξ_i,η_i) and (ξ_i',η_i'), for each object point that is identified in both biplane views. To assess the effects of error in the measured distances D and D′, we introduced positive or negative errors of up to 20 cm in both distances. These simulations included true values of D and D′ equal to 100 cm; Θ_x, Θ_y and Θ_z equal to 70°, − 70° and 70°, respectively; intersection of the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes at (0.5 D, 0.5 D′) with no skew; and ten object points. As shown in Fig. 6, average RMS error in the estimated 3-D object-point configurations did not increase substantially until both D and D′ were in error by approximately 10 cm. We simulated combinations of positive error in D and negative error in D′ also, with similar results. Thus, for the conditions studied, we

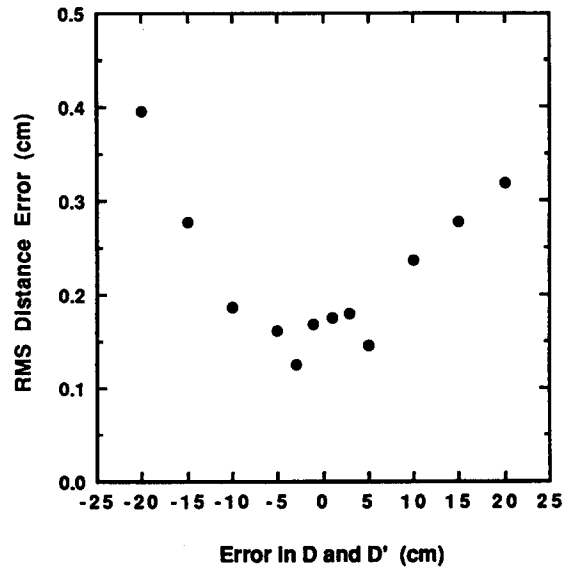


FIG. 6. Dependence of RMS reconstruction error on errors in the measured distances D and D′ for $N = 10$. Errors in D and D′ less than approximately 10 cm do not affect the 3D reconstruction substantially.

conclude that errors of approximately 10 cm in D and D′ are required before RMS object-point error is increased substantially. Since D and D′ are used reciprocally with pixel size to form the scaled image coordinates from the pixel coordinates, the effects of error in measured pixel size may be estimated from the observed effects of error in the measured distances D and D′.

The origins of the image coordinate systems must be determined in order to measure the image coordinates of the object points. Recall that these origins are defined as the position in each image plane where a perpendicular from the corresponding focal spot intersects that plane. Errors in the origin positions were simulated by adding to or subtracting from the true image coordinates of each point a fixed number of pixels in both directions. The biplane geometry was identical to that described above, and again ten object points were included in each configuration. Effects of origin errors of up to 500 pixels in both directions, with positive errors introduced into the coordinates of one image and negative errors into the other, are shown in Fig. 7. It is interesting to note that, in these simulations, errors in origin position greater than 256 pixels corresponded to origin locations completely outside the assumed image-receptor field. From Fig. 7 one can see that errors greater than about 100 pixels were required before RMS error of the calculated object-point locations increased noticeably. Positive errors in the origin positions of both images produced similar 3-D error. These perhaps surprising results can be explained by noting that 100 pixels in the image plane correspond to 3.5 cm and that the distances D and D′ were both 100 cm; thus, a ray extending from the erroneous origin to the focal spot differed from the perpendicular by only 2°.

We examined the effects of error in measured image coordinates by adding uncorrelated normally distributed deviations to the image coordinates of each object point in either one or both views. The resulting image-point

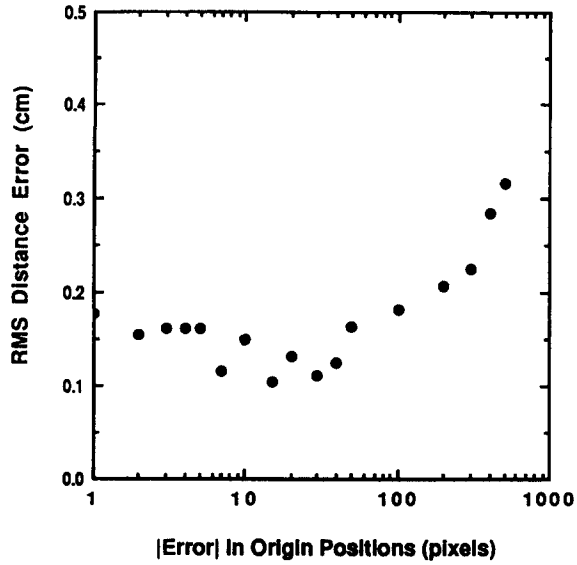


FIG. 7. Dependence of RMS reconstruction error on errors in the origins of the image coordinate systems for $N = 10$. Origin errors on the order of 100 pixels (3.5 cm) or more in both directions, and in both images, were required before reconstruction error increased substantially.

distributions were centered on the exact image location, with standard deviations that varied systematically from 0 pixels to 5 pixels in different simulated experiments. Each image coordinate produced in this way was then rounded to the nearest pixel; thus, an error distribution with a standard deviation of 0 pixels produced error-free image coordinates rounded to the nearest pixel, as in earlier experiments. Figure 8 shows the effects of this image-coordinate error on RMS 3-D object-coordinate error for the situations in which error was added to the coordinates in one or both views. Substantial sensitivity of the 3-D object-point estimates to image-coordinate error is evident in these figures, though the

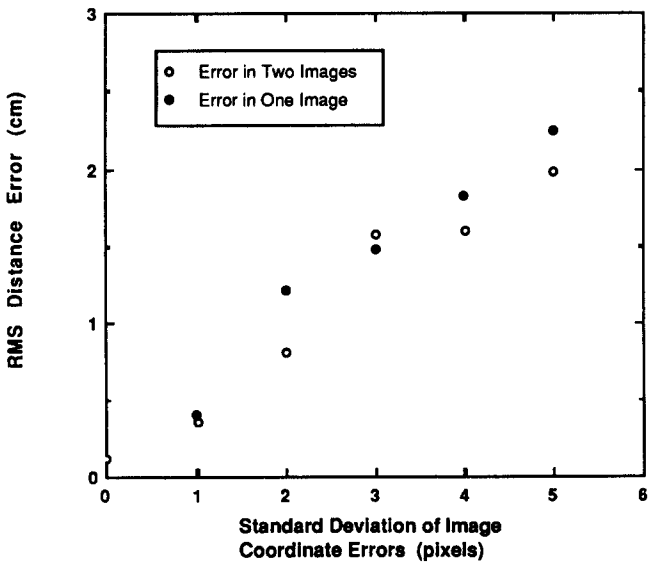


FIG. 8. Effects of normally distributed image-coordinate error in one or both views on RMS reconstruction error, for ten object points.

similarity of the results obtained with image-coordinate errors in one or both views is surprising.

We argued earlier that the condition number of $[A]^T[A]$ should correspond to the accuracy with which we are able to determine $[Q^*]$ and hence the object-point locations, and this relationship was evident in Fig. 3. To study the relationship further, we adopted a single biplane geometry and generated 100 different configurations of ten object points ($N = 10$). Linear regression of average 3-D reconstruction variance against average condition number produced a slope of 0.96 and a correlation coefficient of 0.66. Although this near-unit slope indicates that condition number provided a good indicator of reconstruction error on the average, the low correlation indicates that condition number is clearly not the only factor which determines reconstruction accuracy.

V. DISCUSSION

The computer simulation experiments reported here indicate that the described technique should allow three-dimensional structure of (cardiovascular) objects to be recovered from real biplane images. We have shown that the method performs well for a variety of imaging geometries, and that the accuracy of the 3-D information obtained is not affected strongly by errors in measured quantities other than the image coordinates.

Throughout our simulation experiments, we found the condition numbers of the matrix $[A]^T[A]$ to be quite large, ranging from 10^5 to 10^{12}. Still, the 3-D object-point estimates were generally quite good, with RMS errors commonly in the range of 0.1 to 0.2 cm. It should be noted that we added several steps to the basic reconstruction algorithm described by Longuet-Higgins[11] to improve the robustness of the 3-D information recovered in this inherently ill-conditioned solution task. Among these steps were use of the third unit eigenvector of $[Q^*]^T[Q^*]$ as an estimate of the unit translation vector $\hat{\mathbf{t}}^*$, and calculation of $[Q^{****}]$ via Eq. (5), which guarantees that the matrix $[R]$ estimated from $[Q^{****}]$ and $\hat{\mathbf{t}}^*$ via Eq. (7) will be unitary. For the purpose of comparison, we applied the basic algorithm as proposed originally,[11] without the various additional steps described here and elsewhere,[14] to simulated data in which the image coordinates were known to the nearest pixel and found RMS errors typically an order of magnitude larger than those obtained with our modified algorithm. Thus, the computational steps we have added seem necessary for successful application of the technique to radiographic images.

Figures 2(a) and (b) illustrate the effects of the number of object points on RMS error and condition number when "perfect" data were employed. It is apparent that the *inherent* conditioning of the problem under "perfect" conditions improves dramatically as N increases from eight to nine or ten. It is also apparent from Fig. 2(b) that the effects of computational roundoff error overshadow those of condition number for large values of N. Figure 3 shows that this was not the case when imperfect data were used, however. There, the image coordinates were known only to the nearest pixel, and the resulting variance closely followed condition

number. In this situation, where experimental error was not negligible, condition number rather than computational roundoff error proved to be the primary determinant of accuracy.

Effects of the actual imaging geometry on 3-D reconstruction accuracy were studied thoroughly. The effect of rotation between the views was shown to be minimal when the three rotation angles were between 40° and 140°; thus, it is evident that nearly orthogonal imaging views are not required for accurate recovery of 3-D information with this method.

Other parameters that specify the actual biplane geometry are the distances D and D′, the amount of skew, and the point at which the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes intersect (or approach most closely). We found that RMS reconstruction error varied approximately linearly with D and D′ in the range of values studied, due to the dependence of the angle subtended by the object points on the values of D and D′. From Fig. 1 and the definitions of (ξ_i,η_i) and (ξ_i',η_i'), it is evident that the scaled image coordinates are simply the tangents of the angles between the perpendicular to each image plane (i.e., the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes) and components of the ray that extends from the focal spot to the corresponding image-point location. As D and D′ become larger, the tangents of these angles become smaller in magnitude. Equations (2) and (4) show that as the scaled image coordinates of a particular object point approach zero, all elements except the last in the corresponding row of $[A]$ approach zero as well. Thus, the matrix $[A]^T[A]$ becomes increasingly close to singular, in turn increasing the solution's sensitivity to any error in the input data. This explains not only the dependence of RMS reconstruction error on the values of D and D′, but also the dependence of that error on the amount of skew in the system. As stated above, the common volume of the system, and thus the solid angle subtended by the object points, becomes smaller as skew becomes larger. The relationship between the location at which the $\hat{\mathbf{z}}$ and $\hat{\mathbf{z}}'$ axes intersect and RMS error was pointed out in the previous section.

Overall, we found the method to be remarkably resilient to errors in the measured distances D and D′ and in the origins of the image coordinate systems. The fact that the scaled image coordinates represent tangents of angles explains this happy and perhaps surprising result also.

To investigate the possibility that errors in different measured parameters might reinforce each other, we performed simulations with image coordinates that were known only to the nearest pixel, errors of + 5 and − 5 cm in D and D′, and simultaneous errors of 10, 20, 30 and 40 pixels in the positions of the image coordinate system origins. The resulting RMS reconstruction errors ranged from 0.1 to 0.2 cm, thus ruling out any substantial synergistic effect of combined measurement errors, at least for these error values.

Recall from the previous section that we found reconstruction variance to correlate strongly with condition number. Additional analysis of the same 100 simulations showed little or no correlation between variance and λ_9 alone. Hence, 3-D image reconstruction variance seems to be governed primarily by the condition number of $[A]^T[A]$ when image coordinates are known to the nearest pixel and N is constant. This result contrasts with that found when image coordinates were known to the nearest pixel but the number of object points (N) was varied. Recall that λ_9/N represents the variance of the rows of $[A]$ in the direction of the solution $[Q^*]$. Figure 9 shows that λ_9/N was inversely related to reconstruction error for values of N ranging from 9 to 30, though condition number also was a good predictor of 3-D reconstruction error in these experiments, as indicated by Fig. 3.

Turning now to the effects of error in the measured image coordinates, we see from Fig. 8 that 3-D reconstruction error rose markedly as the standard deviation of the coordinate errors increased from 0 to 5 pixels. Relationships between image-coordinate error and the values of λ_8 and λ_9, for a fixed number of object points, are shown in Fig. 10(a). Not surprisingly, the value of λ_9 increased with image-coordinate error, because the vectors corresponding to the rows of $[A]$ deviated more from the best-fit hyperplane when the errors in the image coordinates increased. The average value of λ_8 (which governed condition number) also increased slightly as more error was introduced, indicating that the rows of $[A]$ became more dispersed within the best-fit hyperplane as well. As shown in Fig. 10(b), λ_9 was a better predictor than condition number of 3-D reconstruction error when image-coordinate error was not constant.

In summary, then, condition number appears to be a good predictor of error in calculated object-point configurations across situations where the magnitude of image-coordinate error is relatively constant. In that case, the ratio λ_9/N by itself does not predict 3-D reconstruction error well. However, in situations where the amount of error in the image coordinates changes but the number of object points (N) does not, the value of λ_9 (or λ_9/N) correlates strongly with the resultant error, and condition number can be quite misleading. These observations carry implications for the eventual implementation of this technique with clinical images,

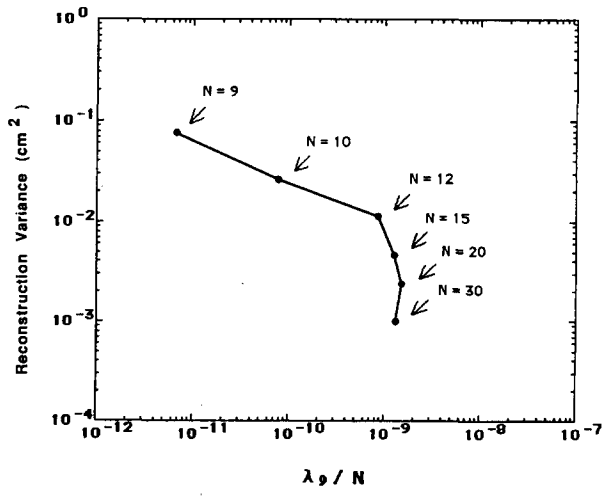
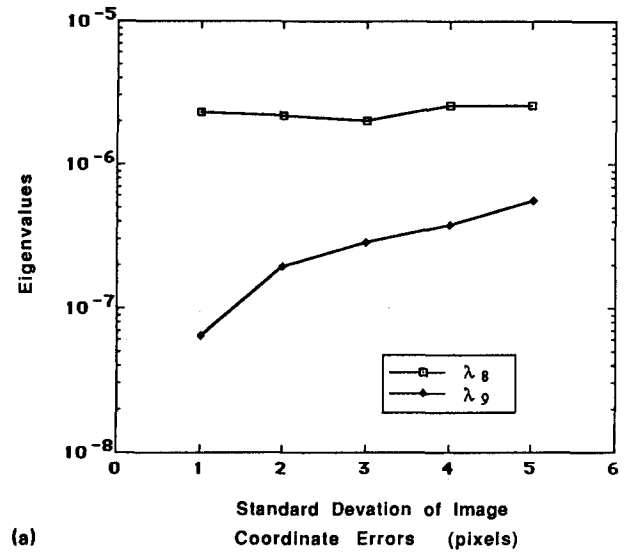


FIG. 9. Relationship between λ_9/N and average mean-square reconstruction error in the calculated object-point configurations for various numbers of object points, with image coordinates measured to the nearest pixel. Each datum represents an average of 50 different object-point configurations.



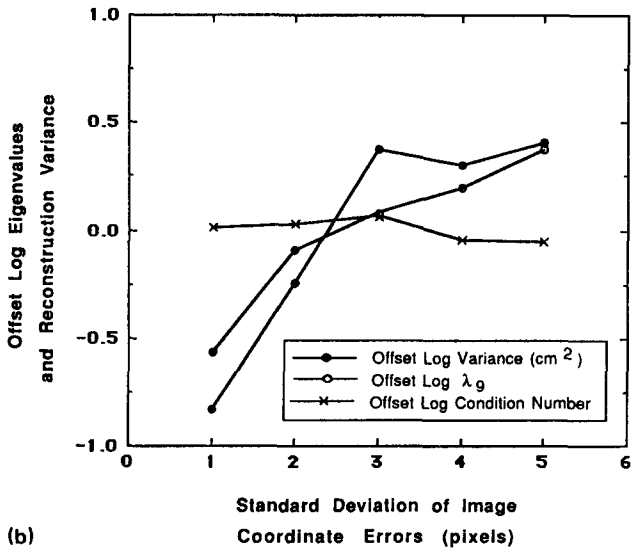


FIG. 10. Effects of errors in the image coordinates of one view, for $N = 10$. (a) Dependence of the two smallest eigenvalues of $[A]^t[A]$ on the standard deviation of image-coordinate error. (b) Dependence of condition number, λ_0 and reconstruction variance (mean-square 3D distance error) on the standard deviation of image-coordinate error. The logarithms of the condition number, λ_0 and variance have been offset by their respective means to facilitate comparison.

where error in the measured image coordinates is likely to constitute the most important limitation of accuracy in the recovery of 3-D information.

As mentioned in Sec. III, we evaluated the precision of the calculated object-point configurations by superimposing each pair of true and calculated configurations as closely as possible. This required a translation of the calculated object points, multiplication of the resulting coordinates by a scale factor, and subsequent rotation of the calculated configuration until the sum of the squared distances between the original and calculated object points was minimized. Values of the required translation fell between zero and tens of cm; values of the scale-factor correction typically were 0.95 to 1.05; and the rotations ranged from fractions of a degree to tens of degrees. In general, the required amount of translation, scaling, and rotation correlated strongly with the minimized RMS error that could be achieved in the 3-D object-point configurations. When large translations and rotations were required, the final accuracy of the calculated configuration usually was poor. Additional investigation will be needed to reveal the factors that cause error in the calculated locations, scales, and orientations of object-point configurations, as opposed to error in the configurations themselves.

VI. CONCLUSIONS

We have shown that the determination of 3-D object structure from biplane images is feasible and potentially accurate without prior knowledge of the relative orientation of the two views. Errors in calculated object-point positions were typically on the order of a few millimeters in our simulation studies, a level of accuracy that may be sufficient in many clinical applications. Experiments that employ both phantom and clinical images are needed now to evaluate the effectiveness of this technique in more realistic situations. Both the speed (on the order of one second at 0.6 MIPS) and the accuracy of the method suggest that it may be useful in clinical settings.

ACKNOWLEDGMENTS

This work was supported by USPHS research Grants CA24806 and CA47043.

[a] Address for correspondence: Box 429, Department of Radiology, The University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637.

[1] K. R. Hoffmann, K. Doi, H-P Chan, and K-G Chua, "Computer reproduction of the vasculature using an automatic tracking method," Proc. SPIE **767**, 449 (1987).

[2] K. R. Hoffmann, K. Doi, H-P Chan, M. Takamiya, "3-D reproduction of coronary vascular trees using the double-square-box method of tracking," Proc. SPIE **914**, 375 (1988).

[3] L. E. Fencil, K. Doi, and K. R. Hoffmann, "Accurate analysis of vessel sizes and stenotic lesions using stereoscopic DSA system," Invest. Radiol. **23**, 33 (1988).

[4] L. E. Fencil, K. Doi, K. G. Chua, K. R. Hoffmann, "Measurement of absolute flow rate in vessels using a stereoscopic DSA system," Phys. Med. Biol. **34**, 1 (1989).

[5] J. H. C. Reiber, J. J. Gerbrands, G. J. Troost, C. J. Kooijman, and C. H. Slump, "3-D reconstruction of coronary arterial segments from two projections," in *Digital Imaging in Cardiovascular Radiology* (Georg Theime, Stuttgart, 1983).

[6] R. O. Kenet, E. M. Herrold, J. P. Hill, J. Waltman, A. Diamond, P. Fenster, J. Barba, M Suardiaz, and J. S. Borer, "Reconstruction of coronary cross-sections from two orthogonal digital angiograms", in *Computers in Cardiology 1986* (IEEE Computer Society, Washington DC, 1987), p. 273.

[7] H. C. Kim, B. G. Min, T. S. Lee, S. J. Lee, C. W. Lee, J. H. Park, and M. C. Han, "Three-dimensional digital subtraction angiography," IEEE Trans. Med. Imaging **MI-1**, 3 (1982).

[8] J. Sitomer, E. G. Anselmo, D. A. Feldt, M. T. LeFree, "One line mathematical model of biplane x-ray gantry," in *Computers in Cardiology 1986* (IEEE Computer Society, Washington DC, 1987), p. 659.

[9] S. A. MacKay, M. J. Potel, and J. M. Rubin, "Graphics methods for three dimensional heart wall motion," Comp. and Biomed. Res. **15**, 455 (1982).

[10] J. M. Rubin and R. E. Sayre, "A computer-aided technique for overlaying cerebral angiograms onto computed tomograms," Invest. Radiol. **13**, 362 (1978).
[11] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," Nature **293**, 133 (1981).
[12] H. C. Longuet-Higgins, "Configurations that defeat the 8-point algorithm," in *Image Understanding 1984*, edited by S. Ullman and W. Richards (Ablex, Norwood, N. J. 1984), p. 173.
[13] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of the three-dimensional motion parameters of rigid objects with curved surfaces," Coordinated Science Laboratory Report R-921 (University of Illinois, Urbana, 1981).
[14] C. E. Metz and L. E. Fencil, "Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: theory," Med. Phys. **16**, 45 (1989).
[15] J. Weng, T. S. Huang, and N. Ahuja, "Motion and structure from two perspective views: algorithms, error analysis, and error estimation," IEEE Trans. Pattern Anal. Machine Intel. **PAMI-11**, 451 (1989).