Determination of three-dimensional positions of known sparse objects from a single projection

Kenneth R. Hoffmann and Jacqueline Esthappan
Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology,
The University of Chicago, Chicago, Illinois 60637

(Received 24 April 1995; accepted for publication 22 January 1997)

A new technique is developed for accurate determination of the three-dimensional position and orientation of known sparse objects, e.g., a configuration of points, from a single-perspective projection. In this technique, a computer model of the known object is translated and rotated so as to align it optimally in a least-squares sense with the projection lines connecting the image points with the focal spot by using a modification of the projection-Procrustes technique. The translational and rotational adjustments are repeated iteratively until the angular change between iterations is less than 0.25°. Simulations indicate that, for rms input image errors of 0.03 cm, the three-dimensional positions and orientations can be determined to within approximately 0.2 cm and 0.3° for a wide range of initially guessed positions and orientations, and positions can be determined with an accuracy of approximately 0.3 cm for objects having as few as four points. In phantom experiments, three-dimensional positions and orientations of a cube phantom were reproducibly determined to within 0.23 cm and 0.13°. The entire calculation requires only 10 s on a VAX 3500 to converge to the solution. The accuracy, precision, and speed of the technique indicate that it will be a useful tool for determination of three-dimensional positions and orientations of known sparse objects.

I. INTRODUCTION

Three-dimensional (3D) analysis is becoming ever more important in the evaluations of patients. Accurate 3D patient positioning is important in the planning and delivery of radiation treatment so that the dose distribution is optimized with respect to the target region and so that other organs and structures are spared (Ref. 1). Localization of implanted electrodes in 3D with respect to the regions of the brain is essential for determination of epileptic foci (Ref. 2). Currently in these two applications, biplane images of the patient are usually acquired, and the known imaging geometry and the determined correspondence between the image data are used for determination of the 3D positions of the structures or objects of interest by applying one of many biplane techniques (Refs. 3–9). Unfortunately, biplane systems are not always available in clinical situations. Thus, 3D positioning techniques which require only single projections may be useful.

Wang and Tsai (Ref. 10) have developed a method for determination of 3D positions of objects based on single projection images for camera geometries, which could be used for x-ray projection geometries. They quote accuracies of 3%–5%. Nguyen and Sklansky (Ref. 11) have proposed methods for determination of the 3D coronary vasculature from single projections. Their technique assumes that the distance of the object from the imaging plane is known approximately, and that the motion of the heart can be described by scaling and rotations. They make use of the motion of the heart to refine their heart model. Although their results are very impressive, their technique may not be readily applicable to stationary objects.

By investigating the implications of the Metz–Fencil and projection-Procrustes techniques, we have determined that for a simple object consisting of a configuration of identifiable points, the projection equations coupled with knowledge of the relative positions of points in the object can be used to determine the 3D location and orientation of that object from a single projection. Simply stated, 3D positions of 3D models, in which the relative positions of points in the model are known, can be determined by properly aligning the model with a single projection, if the correspondence between the object and image points is known. Such a technique has a number of applications, for example, (1) determination of the position of a known 3D vascular tree (precalculated from a previous visit) using a new projection (Ref. 12), (2) determination of the 3D position of a patient [whose 3D anatomy is known from computed tomography (CT) or magnetic resonance (MR) acquisitions] for treatment fractions using single video or x-ray projections, and (3) determination of the 3D position of interventional devices, such as catheters, from single projections, e.g., during fluoroscopy.

Although multiparameter optimization techniques could be brought to bear on this problem (Ref. 13), no method has been presented to date which can solve this problem accurately in a timely manner. The primary reason for this is that the search for the solution must be performed in a six-dimensional space, including three translational directions and three rotational angles. However, by making use of the a priori knowledge of the imaging situation, i.e., that the object must be aligned with its projection lines, the search for the solution becomes simplified. The determination of the translation vector and the rotation matrix can be decoupled for the most part. The refinement of the rotation matrix can be performed by using the projection-Procrustes technique.
By taking advantage of this a priori knowledge, we have been able to develop an accurate and quick method for determining the 3D position of a known object from single-perspective projections.

II. METHODS

A 3D imaging coordinate system, \(xyz\), is defined such that its origin is located at the focal spot used for a particular view, and its axes are specified by the unit vectors \(\hat{x}\), \(\hat{y}\), and \(\hat{z}\). An image plane is located at a distance \(D\) from the focal spot, and vectors \(u\), \(v\), and \(w\) are defined parallel to the vectors, \(\hat{x}\), \(\hat{y}\), and \(\hat{z}\) respectively, with \(u\) and \(v\) lying in the image plane and with \(w\) perpendicular to the image plane. When an image of an object consisting of a number of identifiable points is obtained with this imaging system, the coordinate images \((u_i, v_i)_T\) of the projection of the \(i\)th point in the image plane is related to the 3D position of that point, \((P_i)_T = (x_i, y_i, z_i)_T\) by

\[
(u_i)_T = (x_i)_T \times D/(z_i)_T
\]

and

\[
(v_i)_T = (y_i)_T \times D/(z_i)_T,
\]

where the subscript “\(T\)" implies that these are the “true” positions, as opposed to measured positions which will receive the subscript “\(M\).” The measured coordinates \((u_i)_M\) and \((v_i)_M\) contain errors arising from pixelation, image noise (which can give rise to errors in the calculated image positions), image distortions, such as geometric distortions (Refs. 14 and 15), and the techniques used for determining their location.

Independent of the imaging system, a coordinate system can be defined relative to points in the object itself. The “object” coordinate system is denoted by \(x'y'z'\). The position of the \(i\)th point in the object coordinate system is defined as \((P'_i)_T\) and is related to its position in the “imaging” coordinate system, \((P_i)_T\), by the transformation

\[
P'_i = [R] \times P_i + t,
\]

where \([R]^T\) is a rotation matrix and \(t\) is a translation vector relating the origins of the primed and unprimed coordinate systems. Note that the correspondence between the image points and object points must be known. In this study, the correspondence was determined manually by viewing the calibration object at various orientations.

A linear solution of these equations [Eqs. (1) and (2)] can be obtained which provides an initial estimate of the rotation matrix, \([R]\), and the translation vector, \(t\), relating the imaging and object coordinate systems. These equations can be manipulated to yield equations that are similar to those presented as Eqs. (1a) and (1b) by Metz and Fencil (Ref. 7). The calculation of \([R]\) and \(t\) follows the method derived in their paper, except that in their Eq. (14b), the \(\xi, \eta, \iota\), is replaced, respectively, by \(x', y', z'\) defined in this paper. As in their method, the linear solution results in two rotation matrices and in a unit translation vector, \(t\). The magnitude of the translation vector, \(|t|\), is given by the equation

\[
|t| = \frac{\sum_i (\xi_i \times \hat{t}_i) \times ([R] \times P_i)_T - ([R] \times P_i)_T)}{\sum_i (1 - |\xi_i \times \hat{t}_i|^2)}
\]

where the summations are over all the corresponding points, and \(\xi_i = [u_i/D, v_i/D, 1]^T\). As in the Metz–Fencil technique, there are four possible solutions, arising from the two rotation matrices and the sign ambiguity of \(t\). Two of these solutions yield \(t_z < 0\), which effectively places the object behind the focal spot, and can thus be eliminated. Of the other two that place the object in front of the focal spot, one yields an object that is rotated approximately 180° about \(t\) relative to its true orientation; the other yields the object oriented “correctly” to within the effects of input image error and has the smaller image error (defined below) of these latter two. Thus, by identifying the solution in front of the focal spot with the smaller image error, the solution resulting from the “correct” \(R-t\) combination is obtained. Although this modified Metz–Fencil approach can provide a good first estimate of the orientation and position, we have found that an iterative approach, based on the projection-Procrustes (Ref. 9) technique provides more accurate results (see below).

The iterative approach proceeds as follows. After measurement of \(P'_i\), and the coordinates \((u_i)_M\) and \((v_i)_M\), and by use of the relationships given in Eqs. (1a) and (2), the 3D positions of the object points in the unprimed (imaging) coordinate system are then determined in three steps. (1) An initial positioning of a model of the object and refinement of \(t\) is performed. (2) The orientation of the model is iteratively adjusted using the projection-Procrustes technique (Ref. 9). (3) Refinement of \(t\) is repeated after refinement of the orientation is complete. Steps (2) and (3) are repeated a specified number of times.

A. Initial alignment

Using an initial estimate of \([R]\) and \(t\) (either from the linear solution or from a manually input estimate), the model is translated and rotated according to Eq. (2). The projection of the model is generated by use of the projection Eqs. (1a) and (1b), thereby producing the calculated image data \(u_i\) and \(v_i\). The differences between \((u_i)_M\) and \((v_i)_M\) and \(u_i\) and \(v_i\), respectively, are calculated and averaged, i.e.,

\[
dif u = \Sigma_i ((u_i)_M - u_i)
\]

and

\[
dif v = \Sigma_i ((v_i)_M - v_i),
\]

where the index \(i\) runs over the number of identified points in the image. The quantities \(dif u\) and \(dif v\) indicate the error in the center of mass of the model along the \(x\) and \(y\) axes in the imaging plane, respectively. Next, \(t\) is adjusted on the basis of these differences, according to

\[
t_x = (t_x)_\text{prev} - \text{dif } u^* (t_x)_\text{prev}/D
\]

and

\[
t_y = (t_y)_\text{prev} - \text{dif } v^* (t_y)_\text{prev}/D,
\]

where the subscript “\(\text{prev}\)” indicates components of the previous \(t\). The factor \((t_z)_\text{prev}/D\) appears because the errors
must be corrected by the magnification used in the calculation of \( u \) and \( v \), i.e., by the previous magnification.

For the adjustment of \( t_z \), the projection-Procrustes technique (Ref. 9, summarized briefly below) is applied. The scaling factor relating the points in the model and their projections onto the projection lines is calculated, and \( t_z \) is then adjusted by this scaling factor.

**B. The projection-Procrustes technique**

The projection-Procrustes technique rotates, translates, and scales the model so as to align the points in the model optimally, in a least-squares sense, with their corresponding projection lines. For optimal alignment of the set of points with the projection lines, the sum of the squares of the distances between the reconstructed points and their respective projection lines is minimized.

The distance between a particular point and its respective line is given by

\[
d_i = |\mathbf{P}_i - (\mathbf{P}_i)_p|,
\]

where \((\mathbf{P}_i)_p\) is the projection of \( \mathbf{P}_i \) onto its respective projection line. This projection is given by

\[
(\mathbf{P}_i)_p = \mathbf{P}_i \hat{h}_i,
\]

where \( \hat{h}_i = \left[ (u_i)M, (v_i)M, D \right]/\left[ (u_i)^2 + (v_i)^2 + D^2 \right]^{1/2} \) is the unit vector that represents the direction from the focal point (i.e., the origin of the \( x'yz \) coordinate system) to the \( i \)th image point in the imaging system.

The minimization of \( \Sigma(d_i)^2 \) can be achieved by use of a Procrustes algorithm (Refs. 17–19), which determines the translation, rotation, and scale factor that matches two \( n \)-dimensional configurations of points in the least-squares sense. In our case \( n \) equals 3, as we are dealing with two 3D sets of points, i.e., \( \mathbf{P}_i \) and \((\mathbf{P}_i)_p\). For the Procrustes algorithm, \( \mathbf{P}_i \) and \((\mathbf{P}_i)_p\) are employed as the original and target configurations, respectively.

It should be noted that this application of the projection-Procrustes technique differs from that in the previous technique (Ref. 9) in that the model is not positioned with its center of mass at an arbitrary \( z \) location, i.e., \( D/2 \), because it is positioned based on the calculated magnification of the object and differences between the projections of the model and the actual image data. In addition, the calculated scaling is ignored because the size of the object is known. Thus, the projection-Procrustes technique is used only for determining the rotation required to align the points in the model with their respective projection lines.

**C. Iterative refinement of the alignment**

When the Procrustes technique is applied, the rotation matrix and the translation vector relating \( \mathbf{P}_i \) and \((\mathbf{P}_i)_p\) are calculated. If no rotation is required, i.e., if \( \mathbf{P}_i \) and \((\mathbf{P}_i)_p\) are already optimally aligned, the rotation matrix returned by the Procrustes technique will have zeros in the off-diagonal elements, \( r_{12}, r_{13}, \) and \( r_{23} \). If small rotations are required, these off-diagonal elements represent misalignments and are subsequently used for refinement of the rotations applied to the object.

Starting with the estimate of the position and orientation of the model, an iterative process of refining the rotation matrix and the translation vector is begun. This refinement process proceeds in the following manner. The off-diagonal elements, \( r_{12}, r_{13}, \) and \( r_{23} \), are monitored, and the best estimates of the rotation angles in the \( xy, xz, \) and \( yz \) planes (\( \Theta_{xy}, \Theta_{xz}, \) and \( \Theta_{yz} \), respectively) are stored. These angles are modified by the respective off-diagonal elements one at a time, i.e., first the \( xy \) rotation, etc. The model is then rotated by the rotation matrix generated by use of the modified angle. The projection-Procrustes technique is applied, and the respective off-diagonal element is compared with the previously calculated off-diagonal element. If the projection-Procrustes technique returns an off-diagonal element that is smaller than the previously calculated one, the algorithm assumes that the model is better aligned with the more recent adjustment than without it. The adjusted angle is therefore taken as a better estimate of the rotation in that plane. The new respective off-diagonal element returned by the projection-Procrustes technique is then taken as the adjustment for the next iteration. This process continues for a given rotation axis until the angle returned by the projection-Procrustes technique is either greater than the adjustment angle or less than a predetermined threshold. Each of the three angles is refined individually in this iterative manner. After each angle has been refined in this way, the translation vector is refined as described in Sec. II A, by use of the model rotated by the refined angles. This sequence of steps is repeated 40 times. This number of iterations was chosen because it provided results that were relatively insensitive to the setting of the initial angular threshold [see Sec. III, Fig. 6(b)]. At each of these 40 iterations, the threshold for halting the iteration is reduced, i.e., the threshold is set to (initial angular threshold/\( n \))^2, where the initial angular threshold is the angle [usually set to 10° (see below)] below which the refinement procedure stops for the first iteration, and \( n \) is the number of iterations.

Repetition of these steps was found to be necessary because this process could not yield the desired result with a single iteration due to the coupling of the angles in the rotation matrix as well as their coupling with the estimates of the translation vector. In addition, the coupling of the angles affected the convergence of our approach when a single threshold was used. After one angle had been refined, it would be “readjusted” (often upward) during the refinement of a subsequent angle. By setting an initially rather high angular threshold and reducing it gradually, all angles appear to converge more or less at the same rate, and the upward readjustment is not as substantial because no one angle is refined more than the others. At the final iteration, the threshold angle is set to 0.25°, when the initial angular threshold is set to 10°.

After these iterations, the model is rotated and shifted by the determined rotation matrix and translation vector. The discrepancies between the projected model data and the im-
age data are calculated, and the rotation returned by the projection-Procrustes technique is displayed. The off-
diagonal elements indicate residual errors in the angles of rotation about the respective axes. These are usually of the
order of 0.01°.

There is more than one rotation matrix to which the tech-
nique can converge, because the technique requires only that
the off-diagonal elements of the refinement rotation be approximately 0. Thus, rotation matrices containing two −1’s-along the
diagonal (the determinant must be greater than 0 for right-handed
rotations) can be returned. For these rotation matrices, the
calculated configurations resulting from the inappropriate ro-
tations are rotated 180° about one of the primary axes rela-
tive to the true configuration and yield large image error.
Therefore, when two of the diagonal elements are approxi-
mately equal to −1, 180° is added to the calculated rotation
about the appropriate axis, and the iterative refinement of the
alignment is repeated.

D. Evaluation of the technique

To evaluate the accuracy of the results obtained with this
technique, we performed Monte Carlo simulation studies. In
our simulations, we employed a single-plane geometry. The
distance from the focal spot to the imaging plane was 98 cm.
A model of a cube phantom (10 cm on a side) containing 12
points on its surface was generated such as shown in Fig. 1.
The points on the cube were positioned in three of the cor-
ers, in the centers of three faces, and along the diagonals of
the three other faces. The simulation model was positioned
with point one of the simulation model at (0, 0, 70 cm), (0, 0,
50), and (0, 0, 30) in the imaging system (i.e., magnifications
of approximately 1.3, 1.8, and 2.8, respectively) to evaluate
the effect of magnification. After positioning, the model was
rotated about its origin by an angle of 30° about the y axis.
Other orientations (see Table I) yielded results equivalent to
those reported here to within one standard deviation.

Prior to projection of the model into the imaging plane, a
Gaussian-distributed error with a full-width at half-
maximum (FWHM) of 0.001, 0.05, 0.1, or 0.3 cm was added
independently to the x, y, and z coordinates of the simulation
model points. This added error we call “input object error.” The 3D configuration of points (containing error) was then
projected onto the imaging plane, and the Gaussian-distributed error with a FWHM of 0.003, 0.03, 0.1, or 0.3 cm was also added independently to the x and y co-
ordinates of the image points. This added error we call “in-
put image error.” These levels of input errors (see Table I)
both for the model and for the image data span the range

<table>
<thead>
<tr>
<th>Table I. Summary of the simulations.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model orientation = (0°, 30°, 0°)</td>
</tr>
<tr>
<td>Model position = (0, 0, 50) cm</td>
</tr>
<tr>
<td>Number of points in model = 12</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Input random angular error (maximum error) = 20°</td>
</tr>
<tr>
<td>Initial angular threshold = 10°</td>
</tr>
<tr>
<td>Input object error (FWHM) = 0.01 cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.001 cm</td>
</tr>
<tr>
<td>Model orientation = (0°, 30°, 0°)</td>
</tr>
<tr>
<td>Model position = (0, 0, 50) cm</td>
</tr>
<tr>
<td>Initial angular threshold = 10°</td>
</tr>
<tr>
<td>Input object error (FWHM) = 0.05 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 30), (0, 0, 50), (0, 0, 70) cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 50) cm</td>
</tr>
<tr>
<td>Number of points in model = 12</td>
</tr>
<tr>
<td>Input object error (FWHM) = 0.05 cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 30), (0, 0, 50), (0, 0, 70) cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 50) cm</td>
</tr>
<tr>
<td>Number of points in model = 12</td>
</tr>
<tr>
<td>Input object error (FWHM) = 0.05 cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 30), (0, 0, 50), (0, 0, 70) cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 50) cm</td>
</tr>
<tr>
<td>Number of points in model = 12</td>
</tr>
<tr>
<td>Input object error (FWHM) = 0.05 cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
<tr>
<td>Model position = (0, 0, 30), (0, 0, 50), (0, 0, 70) cm</td>
</tr>
<tr>
<td>Input image error (FWHM) = 0.003 cm</td>
</tr>
<tr>
<td>Number of refinement iterations = 40</td>
</tr>
</tbody>
</table>

Medical Physics, Vol. 24, No. 4, April 1997
which we believe will be encountered in clinical situations.

For each model and image data set, a second model was generated. This second model was then aligned with the projection lines and image data using the technique described above. Its final 3D position and orientation were compared with that of the original model to evaluate the technique. The initial position of point one of this second model was always (0, 0, 70 cm).

To evaluate the sensitivity of the technique to the starting position, the initial orientation of the second model was varied by choosing angles angle randomly between −20°, and 20° for each of the angles of the major axes. In addition, a second set of simulations were performed in which the initial orientations were varied up to 50° for each of these axes.

To evaluate the described techniques in a real imaging situation, radiographic images were obtained of a Plexiglas™ cube phantom containing 12 lead shot beads, each 0.1 cm in diameter (similar to the model used in the simulations, see Fig. 1). The positions of the beads on the cube were measured to the nearest 0.05 cm with a ruler that had 0.1 cm markings. The Plexiglas cube was placed near the isocenter of an x-ray imaging system. Prior to application of the above discussed techniques, the imaging system was calibrated, i.e., source-to-image plane distances were measured, pixel sizes were determined. Because film was used, no distortion correction was employed. The source-to-image-plane distance, D, was measured to be 140 cm (±0.5 cm). Film images of the stationary cube were obtained at 30° right anterior oblique (RAO), 0° RAO, 30° left anterior oblique (LAO), and 60° LAO. These film images were digitized using a laser film digitizer (Lumisys Corp., Sunnyvale, CA) with a spot size and sampling distance of 0.0210 and 0.0174 cm, respectively. After digitization, the image was reduced by a factor of 2 by subsampling. The pixel sizes along the x and y axes were determined to be 0.0351 ± 0.001 and 0.0343 ± 0.001 cm, respectively. The positions of the projections of the lead shot beads in the images were indicated by two separate observers. After indication, the local maximum was identified automatically, the average pixel value in a 5 × 5 pixel region about this maximum was calculated, and the center of mass was calculated for these pixels with pixel values above this average. The average difference between input image data of the two observers was 0.008 ± 0.002 cm, i.e., approximately 0.2 pixel.

To evaluate the accuracy of this technique, indices of error were calculated for each configuration and then were averaged over the configurations. Errors in the orientations of the calculated object, the errors in the 3D positions of the calculated points, and the image error were calculated as follows.

Rather than compare the calculated Θ_{xy}, Θ_{xz}, and Θ_{yz} with the initial angles, the Procrustes algorithm was used to calculate the rotation matrix, [R]_j, required to rotate the calculated configuration of points of simulation j into the true configuration. This rotation was taken to be the “orientational” or angular error of the calculated configuration. The mean angular error was then calculated as

\[
E_R = \frac{1}{M} \sum_{j=1}^{M} \cos^{-1}\left[\frac{\text{Tr}([R]_j) - 1}{2}\right],
\]

where M is the number of simulations. The mean error in the absolute 3D positions of the calculated 3D points (here called “absolute 3D error”) was calculated as the average distance of each reconstructed 3D object-point location from its true position in the unprimed system. The mean error in the calculated image coordinates (here called “image error,” which should not be confused with the input error that had been introduced into the image data) was calculated as the average rms distance between the coordinates of the projections of the reconstructed points onto the imaging plane and the coordinates of the input image data (which contained error).

II. RESULTS

A. Simulations

As discussed above, the rotation matrix can be determined by a linear solution of Eqs. (1) and (2), and |t| can be determined by Eq. (3). These estimates can be used as a starting guess. This approach was employed and compared with the results when a random guessed starting position and orientation were used. The 3D absolute and angular errors resulting from these two approaches were not substantially different, lying within one standard deviation of each other. However, when the linear technique was employed without iterations, the 3D absolute and angular errors were a factor of 3.6±0.5 greater than those obtained using the above described technique after guessed random positions and orientations for ranges of input image and object error as used below. Thus, the iterative technique is approximately 3–4 times more accurate than the linear technique in determination of the position and orientation of the object whether or not the results of the linear technique are used as an initial guess.

The results obtained with a range of ±50° for the initial orientation yielded mean absolute 3D errors as well as mean angular errors which lay well within one standard deviation of those obtained with a 20° range, with the 50° range yielding slightly smaller errors, indicating that the iterative technique does not depend on the initial starting position. The results discussed below were generated using an orientation based on an initial guess angle for each angle, Θ_{xy}, Θ_{xz}, and Θ_{yz}, derived from a random distribution of angles (FWHM = 20°) about the true angles.

We present results illustrating the effect of input image error on the absolute 3D error and the angular error, in Figs. 2(a) and 2(b), respectively. The object error was set to 0.001 and 0.05 cm to observe its effect on the calculated configuration. The center of mass of the model to be aligned with the projection lines was positioned at approximately 75 cm from the focal spot. For small object error (0.001 cm), the mean absolute 3D error and the mean angular error track linearly with the input image error. For a larger object error (0.05 cm), these errors plateau for input image error levels below 0.03 cm, indicating that the object error dominates in this region. The errors converge for input image error levels
above 0.03, indicating that image error dominates in this region. For 0.03 cm input image error and 0.05 cm object error, positions and orientations are accurate to within 0.2 cm and 0.3°, respectively.

We present results illustrating the effect of input object error on the absolute 3D error and the angular error in Figs. 3(a) and 3(b), respectively. Here results are similar to those seen in Fig. 2 in terms of areas where the two types of errors dominate. Although not shown here, the input object error and calculated errors show a linear relationship when the image error is small (0.003 cm). These input image error levels (and thus the results obtained with them) reflect the range of input image errors that we expect in the experimental situation. These results indicate that positional and angular errors will be approximately equal to or less than 1 cm and 1°, respectively.

We present results illustrating the effect of input image error on the absolute 3D error and the angular error [in Figs. 4(a) and 4(b), respectively] for the orientation of the model with different numbers of points on the cube. The input object error was set to 0.05 cm. The dependence on the input image error is again evident here. Although the error bars of each set of data tend to overlap, there appears to be a dependence of the absolute 3D and angular errors on the number of points in the object, with errors increasing as the number of points decreases. The deviation of the results of the eight point configuration from this trend for small input image error may be due to fluctuations in the starting point and/or
quantization of the final angles (i.e., stopping when the angles change by 0.25°). However, even with as few as four points in the configuration, one pixel input image error results in average absolute 3D and angular errors of approximately 0.2 cm and 1°, respectively. Although these results indicate that accurate 3D positions can be determined with as few as four points, we have found that the technique fails when the points lie in a plane, due to degeneracy in the solution.

We present results illustrating the effect of input image error on the absolute 3D error and the angular error [in Figs. 5(a) and 5(b), respectively] for the orientation of the cube at three different magnifications of the cube. The input object error was equal to 0.05 cm. In addition to showing dependence on the input image error, the results now show a dependence on the magnification, with errors decreasing as the magnification increases. This is probably due to a "demagnification" of the input image errors, i.e., the input image error is reduced by the magnification factor at the position of the 3D point.

We investigated the effect of the initial angular threshold used to halt the iterative refinement as well as the effect of the number of iterations used in the refinement process on the accuracy of the orientation of the model, i.e., the angular error. In Fig. 6(a), the angular error is plotted as a function of...
the initial angular threshold for the two cases when 20 and 40 iterations were used in the refinement process. The input image and object errors were 0.03 and 0.05 cm, respectively. When 20 iterations are used, the angular error increases approximately linearly with the value of the initial angular threshold, whereas when 40 iterations are used, the angular error appears to be independent of the value of the initial angular threshold. This is probably due to the fact that when 40 iterations are used, 39 iterations are performed for the same angular range in which 20 are performed when 20 iterations are used, thus, leading to improved and more robust refinement. In Fig. 6(b), we present the dependence of the angular error as a function of the number of iterations used, for the two cases with the initial angular threshold equal to 10° and 20°, with the same input errors as in Fig. 6(a). There is a strong dependence on the number of iterations used in the refinement process, however, above 40 iterations the angular error does not appear to depend on the number of iterations. These results indicate that the technique may yield reliable, accurate, and consistent results when 40 iterations are used in the refinement process, independent of the initial angular threshold used. The absolute 3D error was found to track with the angular error in a manner similar to that seen in Figs. 2–5.

B. Phantom experiments

In phantom experiments, we could not measure the actual positions of the lead beads to the level of precision and accuracy provided by the method presented above. Hence, we chose to rely on image error, the magnitude of the residual errors in the rotation relating the final configuration and the projection lines (projection-Procrustes rotational error), and the consistency in the data calculated from two images acquired at the same angle. The average image error for the combined observer data was 0.038 ± 0.005 (0.092±0.007 cm for the linear technique), which is comparable to that obtained for the 0.03 cm input image error in the simulations. The projection-Procrustes rotational errors were 0.015° ± 0.008° and 0.197° ± 0.100°, for the iterative and linear techniques, respectively. The average distance between configurations calculated from the two sets of observer data was 0.048 ± 0.39 cm. The good agreement between the data of the two observers is due primarily to the robustness and reproducibility of the method used for determining the positions of points in the image (Sec. II D). For the two angles (0° RAO and 30° RAO) at which two acquisitions were made, the differences in the 3D positions were 0.22 ± 0.012 cm and 0.24 ± 0.12 cm for the iterative and linear techniques, respectively, and the angles agreed to within 0.13° ± 0.15° for both techniques. These results indicate that both the linear and nonlinear technique produce highly reproducible results, but that the iterative technique produces configurations which are better aligned (by more than a factor of 3) with the projection lines.

Errors in the positions were not distributed isotropically in the x, y, and z directions. The errors in x and y were approximately equal to the input image error divided by the magnification. The remaining absolute 3D error in the reconstructed 3D object coordinates, which was the dominant component for large input image error, appears to lie primarily in the z direction. These results and the relatively low angular error indicate that large input image errors propagate into magnification errors in the calculated configuration. These were less than 1% (with 0.03 cm input image error) for imaging geometries similar to those investigated here.

IV. DISCUSSION

We have presented a technique for accurate, automatic determination of 3D positions of known sparse objects from single projections which is simple and quick (10 s on a VAXstation 3500), and for which the position and orienta-

![Fig. 6. (a) Angular error plotted as a function of the initial angular threshold for the angular refinement, when 20 (closed circle) and 40 (open square) refinement iterations were employed. The input image and object errors were 0.03 and 0.05 cm, respectively. (b) Angular error plotted as a function of the number of iterations used in the refinement process, for the initial angular threshold equal to 10° (closed circle) and 20° (open square). The error bars represent one standard deviation of the distribution of the errors.](image-url)
tion of the object relative to a properly calibrated imaging system need not be known a priori. Our simulation results indicate that, with this new single projection technique, one can determine 3D positions of object points with an accuracy of approximately 0.2 cm for an rms input image error of approximately 1 pixel (0.03 cm), independent of the estimated starting angle. This iterative technique yields results that are approximately 3.6 times more accurate than those obtained by a linear solution of the problem. Although the accuracy of the technique depends on the number of points used, an accuracy of 0.2 cm in positioning and 1° in orientation is obtained with as few as four points and an input image error of one pixel. The accuracy of input data is the primary determinant of the accuracy in the resulting position and orientation of the calculated configuration. As the magnification increases, the accuracy of the technique improves. This probably results from the effective demagnification of the input image errors with increasing magnification. By choosing 40 or more iterations in the refinement process, the accuracy of the technique does not appear to depend on the initial angular threshold used. Phantom results indicate that an error in input image data of 0.03 cm can be obtained when the described center of mass technique is used. Three-dimensional positions were determined with a precision of approximately 0.24 cm when the gantry was repositioned for the same imaging geometry.

Additional features of the technique were observed during its evaluation. From the phantom experiments, we noticed that, if the correspondence of the points was incorrect, the technique still converged in that the off-diagonal elements of the refinement rotation returned by the projection-Procrustes technique were approximately 0; but, the image error was then 3–10 times larger than that obtained with the correct correspondence. Thus, the technique may be useful for determination of correspondences between object points and points in the image. Errors in the source-to-image distance, \( D \), propagate into positional errors, changing the effective magnification of the object. Errors in the center of the image propagate directly into comparable (demagnified) errors in the \( x \) and \( y \) positions of the calculated configuration. We have used the result of the linear solution (Sec. II) and found that the results are comparable to those obtained using the random guesses (Sec. II D) as a starting point, indicating that either could be used for the initial positioning of the model. We have found empirically that the technique fails when all points lie in a plane, probably due to a degeneracy in the solution. We have not yet investigated whether the technique would fail for various symmetric structures, such as spheres or ellipses.

As stated in Sec. I, this technique would be useful in a number of situations. We have investigated its application for determination of the position of a known 3D vascular tree (precalculated from a previous visit) using a new projection (Ref. 12). For this application, a 3D model of the vascular tree of a patient is obtained, the subsequent 3D positions of points in that vascular tree are determined with this technique by analysis of a third (single) projection after four or more corresponding points have been identified. This inform-

ation is then used to align the entire 3D vascular tree with the current projection. Thus, the relationship between the gantry and the original 3D vascular tree is known, and the gantry can then be rotated so as to obtain the same projection as previously obtained. In addition, the tree can be rotated by the computer, and optimal views (Refs. 20, 21) can be determined by viewing of the rendered vascular tree, and the gantry can be rotated to that angle for the desired acquisition. This may result in a saving of x-ray exposure to the patient and staff x-ray exposure, and of additional contrast dose to the patient.

This technique is also useful for calculation of the rotation matrix and translation vector relating those two imaging systems based on the determined 3D positions of the model relative to two separate imaging systems (Ref. 22). This technique may be useful in situations where the enhanced Metz–Fencil technique (Ref. 9) cannot be used, e.g., less than eight identifiable points available, or does not yield accurate results. We believe that this technique will also be useful for determination of the 3D position of a patient (whose 3D anatomy is known from CT or MR acquisitions) for treatment fractions using single video or x-ray projections, as well as for determination of the 3D position of interventional devices, such as catheters, from single projections, e.g., during fluoroscopy. These applications are under investigation. We believe that this technique may be useful for any application where a 3D model is available and projection images are acquired.

ACKNOWLEDGMENTS

This work was supported by USPHS Grant Nos. CA47043, T32 CA09649, and HL52567. The authors are grateful to Ryan Reft and Shidong Li, Ph.D., for assistance in data acquisition, to Elizabeth Lanzl for editorial assistance, and to Charles E. Metz, Ph.D., for enlightening discussions.

\(^{a}\)Author to whom all correspondence should be addressed; electronic mail: k-hoffmann@uchicago.edu


\(^{8}\)L. E. Fencil and C. E. Metz, “Propagation and reduction of error in


16 C. E. Metz (private communication).


