Registration of 3-D Images Using Weighted Geometrical Features

Calvin R. Maurer, Jr., Student Member, IEEE, Georges B. Aboutanos, Student Member, IEEE, Benoit M. Dawant, Member, IEEE, Robert J. Maciunas, J. Michael Fitzpatrick, Member, IEEE

Abstract—In this paper, we present a weighted geometrical feature (WGF) registration algorithm. Its efficacy is demonstrated by combining points and a surface. The technique is an extension of Besl and McKay’s iterative closest point (ICP) algorithm. We use the WGF algorithm to register X-ray computed tomography (CT) and T2-weighted magnetic resonance (MR) volume head images acquired from eleven patients that underwent craniotomies in a neurosurgical clinical trial. Each patient had five external markers attached to transcutaneous posts screwed into the outer table of the skull. We define registration error as the distance between positions of corresponding markers that are not used for registration. The CT and MR images are registered using fiducial points (marker positions) only, a surface only, and various weighted combinations of points and a surface. The CT surface is derived from contours corresponding to the inner surface of the skull. The MR surface is derived from contours corresponding to the cerebrospinal fluid (CSF)-dura interface. Registration using points and a surface is found to be significantly more accurate than registration using only points or a surface.

I. INTRODUCTION

Registration techniques quantitatively relate the information in one image to information in another image by determining a one-to-one mapping between the points in each image. Registration of multimodal images makes it possible to superimpose features from different imaging studies. For example, skeletal structures and areas of contrast enhancement seen in X-ray computed tomography (CT) images can be overlaid on magnetic resonance (MR) images to estimate the transformation. Registration techniques have recently been extended to relate image space to physical space. Stereotactic surgery and stereotactic radiosurgery require that an image or images be registered with the physical space occupied by the patient during surgery. New interactive, image-guided surgery techniques use image-to-physical space registration to track in real time the changing surgical position on a display of the preoperative image sets of the patient.

Many methods have been used to register medical images. In this paper we are primarily concerned with point-based and surface-based registration methods. We review existing point-based and surface-based methods, and present a hybrid registration technique first proposed in [38] that uses a weighted combination of multiple geometrical feature shapes. The weighted geometrical feature (WGF) registration algorithm is an extension of Besl and McKay’s iterative closest point (ICP) algorithm and is an improvement over an approach proposed independently in [13]. We demonstrate the efficacy of the WGF algorithm by registering CT and MR volume head images using fiducial points (centroids of rigidly attached external markers), a surface, and various weighted combinations of points and a surface. Registration accuracy is calculated as the distance between marker positions not used to estimate the transformation.

II. REGISTRATION ALGORITHM

A. Point-Based Registration

Point-based registration involves the determination of the coordinates of corresponding points in different images (and/or physical space) and the estimation of the geometrical transformation using these corresponding points. The points may be either intrinsic [15], [24] or extrinsic [11], [17], [36], [57]. Intrinsic points are derived from naturally occurring features, e.g., anatomic landmark points. Extrinsic points are derived from artificially applied markers, e.g., tubes containing copper sulfate.

Registrations involving head images of the same patient are typically rigid-body transformations $T(p) = R p + t$, where $R$ is a $3 \times 3$ rotation matrix, $t$ is a $3 \times 1$ translation vector, and $p$ is a $3 \times 1$ position vector. Let $P = \{ p_j \}$ for $j = 1, \ldots, N_p$ be a point set to be registered with another point set $X = \{ x_j \}$ for $j = 1, \ldots, N_x$, where $N = N_p = N_x$ and where each point $p_j$ corresponds to the point $x_j$ with the same index. We wish to find the rigid-body transformation $T$ that minimizes...
the cost function

\[ d(T) = \sqrt{\frac{1}{N} \sum_{j=1}^{N} ||x_j - T(p_j)||^2}. \]  

(1)

This problem was given the name “Orthogonal Procrustes” problem by Hurley and Cattell [26]; it is known as the “absolute orientation” problem in photogrammetry [20]. A closed-form solution was first discovered by Schönemann in 1966 [51]. Many other closed-form solutions have been independently discovered, including the solution of Arun et al. [2] that is based on the singular value decomposition (SVD) of the position vectors in the two spaces.

B. Surface-Based Registration

Surface-based registration involves the determination of corresponding surfaces in different images (and/or physical space) and the estimation of the geometrical transformation using these corresponding structures. The surfaces are generally represented as sets of large numbers of points. Image space surfaces are typically obtained by segmenting contours in contiguous image slices, and most segmentation algorithms generate contours as a sequence of points. A physical space surface map of the skin can, for example, be created by sweeping the skin with a three-dimensional (3-D) spatial digitizer [32] and [46]. Thus we generally have two point sets \( P \) and \( X \), where there is no correspondence between the points \( p_j \) and \( x_j \).

The principal difference between point-based and surface-based registration algorithms is in the availability of point correspondence information. And it is exactly the lack of point correspondence information in the surface-based registration problem that causes surface-based registration algorithms to be based on iterative search. The general approach is to search iteratively for the rigid-body transformation \( T \) that minimizes the cost function

\[ d(T) = \sqrt{\frac{1}{N_p} \sum_{j=1}^{N_p} ||y_j - T(p_j)||^2}. \]  

(2)

where

\[ y_j = C[T(p_j), X] \]  

(3)

is a point on the surface \( X \)—for example, a set of triangles whose vertices are the point set \( X \)—“corresponding” to the point \( p_j \) and \( C \) is a “correspondence” function. The point set \( P \) and the surface \( X \) have been called respectively the “hat” and “head” [9], [45], [46], the “dynamic” and “static” feature sets [61], and the “data” point set and “model” shape [5]. The image that covers the larger volume of the patient, or the image that has the higher resolution if volume coverage is comparable, is normally used to generate the surface \( X \).

Pelizzari, Chen, and colleagues [9], [45], [46] pioneered the idea of using surfaces to register brain images. They calculate \( y_j \) as the intersection with the “head” \( X \) of a line joining the transformed “hat” point \( T(p_j) \) and the centroid of the “head” points \( X \). They search for the transformation \( T \) that minimizes the cost function \( d(T) \) in (2) using a search technique described by Powell [48].

The time required to evaluate the cost function can be significantly reduced by applying a distance transform to the surface \( X \). The distance transform essentially converts a binary surface image into a gray-level image in which all voxels have a value corresponding to the distance to the nearest surface voxel. Thus the cost function is largely precomputed. At each point in the parameter search space, the cost function is calculated as the root mean square (rms) average of the voxels in the distance image that correspond to each of the transformed points \( T(p_j) \). The distance transform effectively calculates the “corresponding” points \( y_j \) as the “closest” points on the surface. However, it should be noted that the distance transform is spatially quantized, and further that integer approximations of Euclidean distance (e.g., the chamfer 3-4-5 algorithm [6]) are frequently used to save computation time. Thus this approach only approximately minimizes the cost function \( d(T) \) in (2). Barrow et al. [3] originally described a technique using a distance transform to fit edge points from two images and efficiently computed the distance transform using a “chamfering” method. Borgefors [7] improved this “chamfer matching” technique and extended it by including a multiresolution approach. Jiang et al. [27] applied this technique to medical images.

C. Besl and McKay’s ICP Registration Algorithm

Besl and McKay [5] presented a general-purpose, representation-independent, shape-based registration technique which they call the “iterative closest point” (ICP) algorithm. This method can be used with a variety of geometrical primitives including point sets, line segment sets, triangle sets (faceted surfaces), and implicit and parametric curves and surfaces. Point sets are registered using any of the various closed-form techniques developed for the Orthogonal Procrustes problem. All other cases are handled by first assigning one shape to be the “data” shape and the other shape to be the “model” shape. The data shape is decomposed into a point set (if it is not already in point set form). Then the data shape is registered to the model shape by iteratively finding model points closest to the data points, computing the transformation that registers the points, and applying that transformation to the data primitives. Thus the ICP registration method defines the “corresponding” point \( y_j \) to be the “closest” point on the surface. An advantage over the distance transform approach is that subvoxel point and surface position information can be used and further that the exact Euclidean distance is minimized. Besl and McKay prove that the ICP registration method converges to a local minimum of the cost function \( d(T) \) in (2). We note that the proof of convergence depends on the correspondence function \( C \) being

The terms “data” and “model” arise from an industrial application: registration of digitized data from unfixed rigid objects obtained using high-accuracy noncontact devices with an idealized geometric (e.g., computer-aided design) model prior to shape inspection (Besl and McKay wrote [5] while working at General Motors Research Laboratories). When registering 3-D medical image volumes using surfaces, we pick the surface from the image that covers the larger volume of the patient and/or has the higher resolution as the model shape.
the closest point operator (see Section II-E). They suggest a repetitive search using multiple initial transformations to solve the local minima problem and suggest guidelines for the number of initial transformations to use. The technique has been noted by several workers in the medical image registration field [12], [39], and recently implemented for image-to-physical space registration [52].

D. WGF Registration Algorithm

This discussion provides us with a framework in which points, curves (e.g., line segment sets), and surfaces (e.g., triangle sets) can be combined in a meaningful way. Since the ICP registration method is a representation-independent technique, it is naturally suited for registration using different geometrical features. We extend it to provide the following WGF registration algorithm.

Let \( \{ \mathcal{P}_i \} \) for \( i = 1, \cdots, N_p \) be a set of "data" shapes to be registered with another set of "model" shapes \( \{ \mathcal{X}_j \} \) for \( j = 1, \cdots, N_m \). Decompose each shape \( \mathcal{P}_i \) into a point set \( \mathcal{P}_i = \{ \mathbf{p}_{ij} \} \) for \( j = 1, \cdots, N_{p_i} \), if it is not already in point set form. The points to be used for triangle sets, for example, are the vertices. Let \( \{ \mathbf{w}_{ij} \} \) be a set of nonnegative weights associated with the points \( \{ \mathbf{p}_{ij} \} \). We wish to find the rigid-body transformation that minimizes the cost function

\[
d(T) = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} \sum_{j=1}^{N_{p_i}} \mathbf{w}_{ij} \| \mathbf{y}_{ij} - T(\mathbf{p}_{ij}) \|^2}
\]  

where

\[
\mathbf{y}_{ij} = \mathcal{C}_i(\mathbf{p}_{ij}, \mathcal{X}_j).
\]

For this algorithm, \( \mathcal{C}_i \) is defined to be the closest point operator, i.e., \( \mathcal{C}_i \) finds the point \( \mathbf{y}_{ij} \) in the shape \( \mathcal{X}_j \) that is closest to the point \( T(\mathbf{p}_{ij}) \). See [5], [44] for formal definitions of the closest point to a given point on various geometrical shape representations. We use a multidimensional binary search tree to speed up the search process (see Appendix A).

Initialize the procedure by setting \( k = 1, \mathbf{p}_{i0} = \mathbf{p}_{ij} \), and \( \mathbf{p}_{i0} = \mathcal{T}_0(\mathbf{p}_{ij0}) \), where \( \mathcal{T}_0 \) is some initial transformation. The algorithm can be repeated using multiple initial transformations to solve the local minima problem. Iteratively apply the following steps, incrementing \( k \) after each loop, until convergence within a tolerance \( \tau \) is achieved.

1) For each shape \( \mathcal{P}_i \), compute the closest points \( \mathbf{y}_{ijk} = \mathcal{C}_i(\mathbf{p}_{ijk}, \mathcal{X}_j) \) for \( j = 1, \cdots, N_{p_j} \).

2) Compute the transformation \( \mathcal{T}_k \) between the points \( \{ \mathbf{p}_{ij0} \} \) and \( \{ \mathbf{y}_{ijk} \} \) using the weights \( \{ \mathbf{w}_{ij} \} \). In this step, the points \( \{ \mathbf{p}_{ij0} \} \) and \( \{ \mathbf{y}_{ijk} \} \) for all shapes are collectively considered as two corresponding point sets and registered using an extension of the closed-form solution to the "Orthogonal Procrustes" problem developed by Arun et al. [2]. The modification allows the points to be weighted and is detailed in Section II-F below. (The transformation is computed using the initial points \( \{ \mathbf{p}_{ij0} \} \) so that the final registration represents the complete transformation.)

3) Apply the transformation to produce registered points \( \mathbf{p}_{ijk,k+1} = \mathcal{T}_k(\mathbf{p}_{ijk}) \).

4) Terminate the iterative loop when \( d(T_k) - d(T_{k+1}) < \tau \), where \( d(T) \) is given by (4).

E. Convergence Theorem

Theorem: The WGF registration algorithm converges monotonically to a local minimum of the cost function \( d(T) \) in (4).

Proof: The proof is an extension of the proof of convergence of the ICP algorithm [5]. In the \( k \)th iteration of the loop in the WGF algorithm, we first compute the closest points \( \mathbf{y}_{ijk} = \mathcal{C}_i(\mathbf{p}_{ijk}, \mathcal{X}_j) \). The weighted mean square error (MSE) of this correspondence is

\[
e_k = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} \sum_{j=1}^{N_{p_i}} \mathbf{w}_{ij} \| \mathbf{y}_{ijk} - \mathbf{p}_{ijk} \|^2}.
\]

Then we compute and apply the transformation \( \mathcal{T}_k \) to produce \( \mathbf{p}_{ijk,k+1} = \mathcal{T}_k(\mathbf{p}_{ijk0}) \). The weighted MSE after registration and transformation is

\[
d(T_k) = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} \sum_{j=1}^{N_{p_i}} \mathbf{w}_{ij} \| \mathbf{y}_{ijk} - \mathbf{p}_{ijk,k+1} \|^2}.
\]

Since the registration process minimizes the weighted MSE (see Section II-F), \( d(T_k) \leq e_k \). In the \( (k+1) \)th iteration, we compute the closest points \( \mathbf{y}_{ijk,k+1} = \mathcal{C}_i(\mathbf{p}_{ijk,k+1}, \mathcal{X}_j) \). The closest point operator by definition reduces the distance for each point, i.e.,

\[
\| \mathbf{y}_{ijk,k+1} - \mathbf{p}_{ijk,k+1} \| \leq \| \mathbf{y}_{ijk} - \mathbf{p}_{ijk,k+1} \|.
\]

Thus \( e_{k+1} \leq d(T_k) \). Again, since registration minimizes the weighted MSE, \( d(T_{k+1}) \leq e_{k+1} \). Since MSE cannot be negative, we have

\[
0 \leq d(T_{k+1}) \leq e_{k+1} \leq d(T_k) \leq e_k
\]

for all \( k \). Because \( d(T_k) \) is nonincreasing and is bounded below, the WGF algorithm must converge monotonically to a local minimum of the cost function \( d(T) \) in (4).

F. Weighted Point-Based Registration

Let \( \mathcal{P} = \{ \mathbf{p}_j \} \) for \( j = 1, \cdots, N \) be a point set to be registered with another point set \( \mathcal{X} = \{ \mathbf{x}_j \} \) for \( j = 1, \cdots, N \), where each point \( \mathbf{p}_j \) corresponds to the point \( \mathbf{x}_j \) with the same index. Let \( W = \{ \mathbf{w}_j \} \) for \( j = 1, \cdots, N \) be a set of nonnegative weights (see Sections II-D and III-E). We wish to find the rigid-body transformation \( \mathcal{T} \) that minimizes the cost function

\[
d(T) = \sqrt{\sum_{j=1}^{N} \mathbf{w}_j \| \mathbf{x}_j - T(\mathbf{p}_j) \|^2}.
\]

Arun et al. [2] presented a closed-form solution for the special case \( \mathbf{w}_j = 1/N \) that is based on the SVD of the covariance matrix of the position vectors in the two spaces. We find
the rigid-body transformation that minimizes (10) using the following extension of their solution.

1) Translate the point sets \( P \) and \( X \) such that their new origins are their previous weighted centroids \( \bar{p} \) and \( \bar{x} \) and denote the translated point sets and points with a prime, i.e.,

\[
p_j' = p_j - \bar{p} \\
x_j' = x_j - \bar{x}
\]

(11)

where

\[
\bar{p} = \frac{\sum_{j=1}^{N} w_j p_j}{\sum_{j=1}^{N} w_j} \\
\bar{x} = \frac{\sum_{j=1}^{N} w_j x_j}{\sum_{j=1}^{N} w_j}
\]

(12)

2) Calculate the 3 × 3 weighted covariance matrix of the position vectors

\[
H = \sum_{j=1}^{N} w_j p_j' x_j''
\]

(13)

where the superscript \( t \) denotes matrix transposition.

3) Calculate the SVD of the weighted covariance matrix

\[
H = U \Sigma V^t.
\]

(14)

4) Calculate the rotation matrix

\[
R = UV^t.
\]

(15)

If \( \det R = +1 \), then \( R \) is the rotation matrix. If \( \det R = -1 \), then the solution is a reflection rather than a rotation. In this case, the correct rotation is \( R = V'U' \), where \( V' \) is obtained from \( V \) by changing the sign of the column of \( V \) corresponding to the most singular value of \( H \).

5) Calculate the translation vector

\[
t = \bar{x} - R \bar{p}.
\]

(16)

A proof that this algorithm produces the rigid-body transformation that minimizes the cost function in (10) is presented in Appendix B.

III. METHODS

We register CT and MR images using four fiducial points (localized positions of the rigidly attached external markers), one surface (triangle sets constructed from localized contours), and various combinations of these points and the surface using the WGF registration algorithm outlined in Section II-D. We pick the MR points and surface as the “data” shapes and the CT points and surface as the “model” shapes.

A. Image Acquisition

We used X-ray CT and MR volume head images acquired from eleven patients that underwent craniotomies in a stereotactic neurosurgical clinical trial. Each patient had five external markers attached to posts screwed into the outer table of the skull. The locations of the markers were chosen according to individual clinical circumstances and surgeon preference. Each patient also had a COMPASS stereotactic head frame (Stereotactic Medical Systems, Inc., Rochester, MN) applied. The frame was attached to the scanner table during image acquisition. The frame is a redundant reference system that is not used for registration. Nonetheless, it also serves as a head fixation device. The CT images were acquired using a Siemens Somatom Plus scanner. Each CT image contains between 35 and 41 slices that are 3-mm thick (the slice thickness and table advance are both 3 mm, i.e., there is no interslice gap or slice overlap); each slice contains 512 × 512 pixels of size 0.65 × 0.65 mm. The MR images were acquired using a Siemens SP 1.5-Tesla scanner. Images were obtained using the body coil because the stereotactic frame will not fit within the head coil. Each MR image contains 26 slices that are 4-mm thick (with no interslice gap); each slice contains 256 × 256 pixels of size 1.25 × 1.25 mm. Transverse T2-weighted Spin-Echo MR images were acquired for each patient. The imaging parameters are: TE = 90 ms, TR = 3000 ms, two acquisitions, half Fourier reconstruction, slice selection gradient magnitude = 4.8 mT/m, frequency encoding (readout) gradient magnitude = 1.47 mT/m. The readout gradient is oriented in the anterior-posterior direction (vertical in the image slices) with phase encoding in the lateral direction (horizontal in the image slices). An additional MR image was acquired with the identical imaging parameters except that the readout and preparation gradients were reversed.

B. Geometrical Distortion Correction

We have shown that correction of geometrical distortion in MR images can significantly improve the accuracy of both point-based and surface-based registration [37]. We correct MR images for static field inhomogeneity by using the image rectification technique of Chang and Fitzpatrick [8]. A new image, without inhomogeneity distortion, is generated from a pair of distorted images acquired with reversed readout gradients.

C. The Points

The points are localized positions of rigidly attached external markers. The markers are designed to be bright in both CT and MR. They are constructed from hollow plastic cylinders with an inside diameter of 7 mm and an inside height of 5 mm. Plastic marker bases or posts are screwed into the outer table of the skull of the patient. The markers are attached to the posts just prior to image acquisition. Additional detail about the markers, including pictures and image slices showing the typical appearance of the markers in CT and MR images, can be found in [37], [38], and [60].

We define the position of a marker as its centroid and call the determination of this position marker localization. We
calculate an intensity-weighted centroid for each marker using the localization technique described in [60].

We have previously estimated marker localization error (MLE)\(^2\) with clinical trial data [41], phantom experiments [42], and numerical simulations [59]. For the marker shape and size, voxel dimensions, and SNR in our study, the predicted MLE is approximately 0.2 mm. The MLE estimated using phantom experiments and clinical trial data is approximately 0.4 mm. The MLE estimated with phantom data is a true measure of accuracy (as opposed to reproducibility) since these experiments registered localized image positions to physically known positions. The experimentally estimated MLE is slightly higher than that predicted by numerical simulations, probably because of imperfect marker segmentation.

D. The Surface

We localize contours using a semiautomatic method. We find CT image contours corresponding to the inner surface of the skull and T2-weighted MR image contours corresponding to the CSF-dura interface. We use only transverse slices that lie between the top of the head and just above the eyes. The CSF-dura interface closely approximates the inner surface of the skull in the upper part of the head. A notable exception is the superior sagittal sinus.

A number of segmentation methods have been reported in the literature to detect contours in tomographic brain images. These methods range from manual contour tracing to fully automatic contour detection [1], [14], [28], [29], [33]. Careful manual tracing is difficult, tedious, and time consuming. Automatic methods often work on some images but not on others and may sacrifice accuracy for ease of use. We use a semiautomatic method that allows the user to interact with the algorithm to insure that the contours are visually accurate. The user modifies inaccurate contours not by manually editing the contours but by changing values of parameters used in the algorithm. Our semiautomatic segmentation method, which can be used on both CT and MR images despite their different image characteristics, is a type of deformable model [23], [31], [34], [53]. These methods typically involve the minimization of a cost or energy function defined in terms of both contour properties and image characteristics. Our technique is a variation of a method originally proposed by Gerbrands [19]. It consists of a geometrical transformation followed by the minimization of a cost function. The geometrical transformation maps the two-dimensional (2-D) contour problem into a one-dimensional (1-D) minimization problem that can be efficiently solved with dynamic programming techniques.

The geometrical transform we use assumes \textit{a priori} knowledge of the approximate shape of the contour. This initial shape is deformed to fit the current image. In this study, we use a polygonal approximation to the contour and a transform that computes perpendiculars to the edges of this polygon. The transformed image is a rectangular matrix in which each row corresponds to a different position along the initial contour and each column corresponds to bilinearly interpolated intensity values along the perpendicular lines. In order to treat the problem as a 1-D search problem, we assume that each row in the transformed image contains only one boundary point. If there are multiple regions and thus multiple contours in a single image slice (e.g., there are multiple brain regions in inferior transverse slices of head image volumes), each contour is treated as a separate problem. A cost matrix is associated with the transformed image and a minimum cost path is computed in this matrix using 1-D dynamic programming. This minimum cost path is mapped back onto the original image. In this work we perform two iterations of this process with slightly different parameter values for each iteration (this is described in further detail below). To reduce computation time, the transformation is performed in a limited window around the original contour. This window width is an adjustable algorithm parameter. In order to guarantee closed contours, we constrain the minimum cost path by requiring it to begin and end in the same column.

Fig. 1 illustrates our method on a representative MR image slice. The top and bottom rows illustrate the first and second iterations, respectively. The first iteration uses a very crude user-specified initial polygonal contour and is performed on a blurred image to help avoid local minima caused by noise and small objects near the contour such as bone marrow. The resulting contour, which is shown on the blurred image [Fig. 1(d)] and is also superimposed on the original image [Fig. 1(e)], is a close approximation to the actual boundary but generally does not meet our criterion of visual accuracy as is the case here. The second iteration uses the contour from the first iteration as the initial polygonal contour and is performed on the original image. The resulting contour [Fig. 1(h)] is used to create the surface used for registration. The window size is large in the first iteration to allow the user to specify a very crude initial contour. In the second iteration, the window size is narrowed to constrain the final search because the initial contour (the contour produced by the first iteration) is a close approximation to the actual boundary. The search window size is not a critical parameter but it needs to be large enough to include the actual boundary.

In this study we use a four-term cost function

\[
E(i, j) = \sum_{k=1}^{4} \lambda_k E_k(i, j). \tag{17}
\]

The weights \(\lambda_k\) determine the relative importance of the various terms \(E_k(i, j)\) and are adjusted for each type of image.

1) Intensity Term: This term is used to attract the curve to regions of image intensity \(I_0\). It has the form

\[
E_i(i, j) = [I(i, j) - I_0]^2 \tag{18}
\]

where \(I(i, j)\) is the intensity value of the transformed image at row \(i\) and column \(j\). Each row \(i\) corresponds to a different position along the initial contour. Each column \(j\) corresponds to a different position along perpendiculars to the initial contour. The parameter \(I_0\) is adjusted for each type of image.

---

\(^2\)When the markers are used as fiducials, we call this error \textit{fiducial localization error} (FLE). When the markers are used as targets, we call this error \textit{target localization error} (TLE).
2) **Gradient Term:** $E_2$ is a gradient image calculated by applying the 1-D operator

$$E_2(i, j) = I(i, j + 1) - I(i, j - 1)$$

(19)

to the rows of the transformed image. This is equivalent to computing gradients along the perpendiculars to the initial contour. A positive weight $\lambda_2$ is used in MR to attract the curve to light-to-dark (inside-to-outside) intensity transitions. Conversely, a negative weight is used in CT.

3) **Smoothness Term:** For a contour to be smooth, points on consecutive perpendiculars need to be close to each other. We express this constraint as

$$E_3(i, j) = (j_i - j_{i-1})^2$$

(20)

where $j_i$ and $j_{i-1}$ are the column indexes of contour pixels in rows $i$ and $i - 1$, respectively.

4) **Inflation/Deflation Term:** This term controls inflation or deflation of the final contour. It is expressed as

$$E_4(i, j) = -j^2$$

(21)

where $j$ is the pixel's column index in the transformed image. With a positive weight $\lambda_4$, this term will favor pixels toward the outside of the window on which the perpendicular transform has been computed. Contours will thus have a tendency to inflate (expand). Similarly, with a negative weight, contours will tend to deflate (contract).

All of the cost function terms except the smoothness term are "static," i.e., they can be precomputed from the transformed image prior to optimization. The smoothness term is computed during the optimization process.

Thus the two types of input required by the contour localization algorithm are an initial polygonal contour and a set of parameters used to compute the cost function. To reduce the amount of interactive involvement, we draw initial polygonal contours on only a few slices per image volume (six contours per volume with 13 vertices per contour for CT volumes; five contours per volume with eleven vertices per contour for MR volumes). Initial contours on intermediate slices are obtained by linearly interpolating the polygonal vertex positions. We use two sets of empirically determined parameters—one for CT and one for MR—that provide satisfactory results in most cases. Contour accuracy is visually assessed. Minor interactive parameter adjustments are sometimes necessary. When such
adjustments are necessary, we generally use a single set of parameters for an entire image volume. In one patient whose skull surface is discontinuous due to a previous craniotomy, we used a set of parameters in several CT image slices that were different from the parameters used in the rest of the image volume. In slices where the skull is discontinuous, the contour smoothness weight \( \lambda_3 \) was substantially increased.

The first iteration is performed with a search width of 40 and 25 pixels for CT and MR, respectively. In the second iteration the search window is reduced to 20 and 11 pixels, respectively. The window sizes in CT and MR are different because the pixel sizes in CT and MR are different. The actual window sizes are quite similar (approximately 30 mm in the first iteration, 15 mm in the second iteration). The first iteration for MR images is performed after the MR image is blurred using a 5 x 5 pixel uniform (box) filter. This choice is somewhat heuristic but not critical. A similar amount of blurring could be obtained using a Gaussian filter with \( \sigma \approx 1.5 \) pixels.

The stack of contours localized in an image volume is converted into a triangle set using the simple but effective triangulation method of Christiansen and Sederberg [10]. Briefly, triangle set construction is carried on between pairs of adjacent contours. The algorithm works best if contour pairs are mutually centered, of similar size, and of similar shape. The first two requirements are met by mapping the contours onto a unit square centered at \((0, 0)\). The last requirement is generally true for brain contours in the upper part of the head. The process commences by defining the base of the initial triangle as the line segment formed from the first nodes in each contour. To insure that these nodes are proximate, we arbitrarily choose the point in each contour closest to \((0.5, 0)\) as the first node of each contour loop. It then selects one of the next nodes in the two contours as the third vertex of the triangle. The node is chosen that provides the shortest diagonal. This diagonal forms the base of the next triangle, and the process continues until all contour points are included in a triangle.

### E. The Weights

The weights \( w_{ij} \) are calculated as follows. Let \( \{u_i\} \) for \( i = 1, \ldots, N_s \) be a nonnegative set of inter-shape weights (see Section IV for a discussion of inter-shape weight selection). Let \( \{v_{ij}\} \) for \( j = 1, \ldots, N_p \) be nonnegative normalized sets of intra-shape weights. We calculate the weights \( w_{ij} = u_i v_{ij} \). If \( P_i \) is a point set, we use \( v_{ij} = 1/N_p \), i.e., all points are weighted equally. If \( P_i \) is a triangle set, it is decomposed into a point set \( \{p_{ij}\} \) by using the vertices as the points. We calculate \( v_{ij} \) as the sum of the areas of all triangles of which \( p_{ij} \) is a vertex, and then normalize \( v_{ij} \) for that \( i \) by dividing by \( \sum v_{ij} \). Triangles nearer the top of the head are smaller because of the fixed number of points per contour. With this weighting strategy the influence of each triangle is approximately proportional to its size.

Surface points that are incorrectly segmented or that lie on a nonoverlapping surface can adversely affect registration accuracy. We cope with the outlier problem as follows. We let the WGF algorithm iterate until it converges within a tolerance \( \tau = 5 \times 10^{-5} \) mm. We arbitrarily define outliers as points whose distance from the closest corresponding surface after this first pass is greater than the mean distance plus two standard deviations. We set the intrashape weight \( v_{ij} \) corresponding to each outlier equal to zero, renormalize the intrashape weight sets \( \{v_{ij}\} \) for each \( i \), recompute the weights \( w_{ij} \), and reapply the WGF algorithm.

### F. Registration Error Measurement

One measure of registration accuracy is the root-mean-square value of the cost function after registration. We call the value of the point-based registration cost function [see (1)] fiducial registration error (FRE). Similarly, we call the value of the surface-based registration cost function [see (2)] surface registration error (SRE). A more objective measure of registration accuracy is the distance between corresponding points other than those used to estimate the transformation parameters. Because such points might represent surgically targeted lesions, we call such points targets and the corresponding accuracy measure target registration error (TRE). When we use the term "registration error," without a modifier, we mean TRE.

We randomly assign one marker on each patient as a target and the four remaining markers as fiducials. The fiducials are randomly numbered 1 through 4. We use one combination for each type of registration. For example, when we register images using one point plus the surface, we use fiducial 1; when we register using three points, with or without the surface, we use fiducials 1–3. For every type of registration there is exactly one TRE value per patient. TRE is computed as the distance between the target marker positions in CT and MR after registration.

### IV. RESULTS

We localized contours containing 400 points per slice in 25.7 ± 1.3 (mean ± SD; range = 22–30) slices for CT and 300 points per slice in 18.5 ± 0.9 (15–20) slices for MR. We localized more points in CT than MR because CT has a higher resolution. This resulted in a total of 10 280 ± 520 (mean ± SD) points for CT and 5550 ± 270 points for MR.

We investigated the effect of the intra-shape weighting strategy for triangle sets described in Section III-E on CT–MR surface-based registration accuracy. The accuracy obtained by using this weighting strategy, i.e., by using for each triangle vertex a weight that is proportional to the sum of the areas of the triangles of which the vertex is a member (TRE mean ± SD = 0.71 ± 0.33 mm), is slightly but not significantly better than that obtained without using this strategy, i.e., by weighting all triangle vertices equally (0.77 ± 0.34 mm). Although the improvement in accuracy obtained by using this intrashape surface weighting strategy is slight and not significant, we used it when performing all of the registrations described below.

We investigated the effect of reducing the number of surface points on CT–MR surface-based registration accuracy.

3 Two-tailed paired t-test, \( P = 0.05, n = 11 \) (one target for each of the 11 patients).
by uniformly subsampling the contours. The mean SRE and TRE values resulting from registrations performed using various combinations of CT and MR contour subsamplings are illustrated in Fig. 2. Whereas the value of the surface-based registration cost function (SRE) is relatively independent of the number of MR points for a fixed number of CT points, surface-based registration accuracy (TRE) substantially degrades as either the number of MR points or the number of CT points used for registration drops below 100 points per slice. Registration accuracy obtained using 100 CT and 100 MR points per slice (TRE mean ± SD = 0.73 ± 0.33 mm) is not significantly different from that obtained using 400 CT and 300 MR points per slice (0.71 ± 0.33 mm). Since the time necessary to compute a registration increases with the number of points used, we performed all remaining registrations with 100 points per slice (2570 ± 130 total points) for CT and 100 points per slice (1850 ± 90 total points) for MR.

We registered CT and MR images using combinations of four points and one surface using a variety of inter-shape weights. The mean TRE values resulting from these registrations are illustrated in Fig. 3. In this figure, a normalized point set weight of 1.0 means that only points were used for registration, a weight of 0.0 means that only the surface was used, and any other value of the weight means that a combination of points and the surface was used. Several of the interesting results are listed in Table I. As expected, registration using four points only (TRE mean ± SD = 0.72 ± 0.34 mm) is significantly more accurate than registration using three points only (0.87 ± 0.51 mm). Registration accuracy obtained using the surface only (0.73 ± 0.33 mm) is significantly better than that obtained using three points only; it is not significantly different than that obtained using four points only. Registration using a combination of points and the surface was more accurate than registration using only points or the surface. The best registration accuracy was obtained using four points plus the surface (0.59 ± 0.21 mm). This is significantly better than the accuracy obtained using four points only or the surface only.

Table I also lists "99% TRE" values. These values are calculated as mean +2.326 SD by assuming that the TRE values are normally distributed. The 99% TRE values are an estimate of the worst TRE that will be observed in 99 out of 100 registrations.

We visually examined the registrations by reformatting each CT image volume to MR image coordinate space, drawing contours in several MR image slices, and superimposing these contours on the corresponding CT image slices (see Fig. 4). We also created "fusions" of CT and MR image slices by superimposing alternating bands of the two images. These visual inspections revealed that the in-slice registration accuracy is better than one pixel (1.25 mm), which helps confirm the excellent quantitative results reported above.

A commonly reported problem with surface-based registration is the presence of local minima encountered during the optimization search. It is possible that, since we used image volumes acquired with the patient locked to the scanner table by a stereotactic frame, the initial rotation matrix (that is, the identity matrix) was a very good initial starting point, and
Fig. 4. Visual examination of image registration. (a) is an MR image slice with a localized contour corresponding to the CSF-dura interface. (b) shows the contour overlaid on the corresponding CT image slice after reformatting the CT image volume to the MR image coordinate space using the rigid-body transformation estimated with four points and the surface. (c) is a "fusion" of the CT and MR image slices created by superimposing alternating bands of the two images. Registration accuracy can be visually inspected at the image interfaces.

V. DISCUSSION

Most previously reported registration techniques that align 3-D image volumes by matching geometrical features such as points, curves, or surfaces use a single type of feature. Two groups have developed techniques that use multiple features sequentially [18], [30]. Both groups divide the 3-D image registration problem into a pair of 2-D problems. First, the interhemispheric fissures in the two head image volumes are identified and aligned. Then the sagittal slices in one image are two-dimensionally registered to those in the other image, either by interactively translating and rotating one image until contours in that image visually fit the other image [30], or by matching homologous points identified in corresponding sagittal slices [18].

In this paper, we present a technique that uses multiple features simultaneously. The WGF registration algorithm inherits from Besl and McKay's ICP algorithm [5] the capability of using points, curves (e.g., line segment sets), and surfaces (e.g., triangle sets). The WGF algorithm finds the rigid-body transformation that minimizes the cost function in (4). Collignon et al. [13] presented the first registration technique we are aware of that uses multiple features simultaneously. Our technique is similar to their technique, but with several key differences. First, their cost function explicitly lists terms for points and surfaces only, whereas our cost function implicitly allows the use of any type of geometrical feature, including curves. Second, their technique uses Powell's optimization method [48] to find the transformation that minimizes the cost function, whereas our technique uses an extension of the ICP algorithm [5]. Finally, their cost function, when rewritten in the form of (4), has the summation over shapes \( \sum_{i=1}^{N} \) placed outside the square root symbol. We note that our extension of the ICP algorithm requires that the summation over shapes be placed inside the square root symbol (see proof in Section II-E). Meyer et al. [43] recently reported another technique that uses multiple features simultaneously. There are two important differences between their technique and ours. First, their method uses points, lines, and planes and thus is less general than ours is, since lines are special cases of curves and planes are special cases of surfaces. Second, for the case of rigid-

Thus we avoided the local minimum problem. To test this possibility, we arbitrarily rotated and translated the surface contours of several image volumes. We have found in all cases tested so far that if the initial rotation is less than 30°, the final translation varies by a maximum of several tenths of a millimeter and the final rotation varies by a maximum of several tenths of a degree from the results obtained with no initial rotation. The largest difference we observed between the TRE's of an individual target obtained with and without an initial rotation of less than 30° is 0.2 mm; 90% of the TRE differences were less than 0.05 mm. When the initial rotation was larger than 30°, the algorithm often stopped in local minima far from the global minimum. In summary, the local minima we encountered during our searches were either negligibly different from or very far from the global minimum.

We performed all computations on a Sun Microsystems SparcStation 20. Marker localization took several seconds per marker. Contour localization took several seconds per slice. Registration using points only took less than one second. Registration using the surface only or points plus the surface took several minutes (using 100 surface points per slice for CT and MR).

### Table I

<table>
<thead>
<tr>
<th>Registration type</th>
<th>TRE (mm)</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 points only</td>
<td>0.87 ± 0.51</td>
<td>2.06</td>
</tr>
<tr>
<td>4 points only</td>
<td>0.72 ± 0.34</td>
<td>1.51</td>
</tr>
<tr>
<td>Surface only</td>
<td>0.73 ± 0.33</td>
<td>1.50</td>
</tr>
<tr>
<td>1 point &amp; surface</td>
<td>0.68 ± 0.27</td>
<td>1.31</td>
</tr>
<tr>
<td>2 points &amp; surface</td>
<td>0.64 ± 0.24</td>
<td>1.20</td>
</tr>
<tr>
<td>3 points &amp; surface</td>
<td>0.61 ± 0.22</td>
<td>1.12</td>
</tr>
<tr>
<td>4 points &amp; surface</td>
<td>0.59 ± 0.21</td>
<td>1.08</td>
</tr>
</tbody>
</table>

The accuracy of CT-MR registration performed using points and/or the surface. For the registrations performed using a combination of points and the surface, the TRE values correspond to the minima of the four curves in Fig. 3, and the values of the normalized point set weights \( u_i \) at these minima are listed in parentheses. The 99% TRE values are calculated as mean ±2.326 SD by assuming that the TRE values are normally distributed.
body transformation, their method reduces a line (plane) to two points (one point) in each image space before registration. These points are considered as corresponding points and the registration transformation is therefore computed in one step. Our method decomposes a nonpoint shape into a point set (e.g., the vertices if the shape is a surface represented as a triangle set) in one of the image spaces. For each of these points the corresponding point in the second space is then determined at each step of the iterative search.

It is intuitively clear that a combination of geometrical features might improve registration accuracy. But it is also intuitively clear that a combination of features might merely produce a registration with intermediate accuracy. This work quantitatively demonstrates that registration of 3-D head CT and MR image volumes using points and a surface can be significantly more accurate than registration using only points or a surface. However, we note that the registration accuracies obtained using four points or the surface only are nearly identical and those obtained using three points or the surface only differ by less than 20%. It is unclear whether or not registrations obtained using combinations of shapes can be more accurate than registrations obtained using single shapes if the accuracies obtained using single shapes are substantially different, as might be the case with many points vs. a surface.

The cost function using points plus the surface is different from the cost function using the surface only. The cost function using a combination of points plus the surface has a global minimum that is more accurate than the global minimum of the cost function using a single shape (points or the surface). To demonstrate this, we used the transformation produced by registration using four points plus the surface as the initial transformation for registration using the surface only. During the iterative search process, the transformation moved away from this initial transformation. The final TRE (mean of 11 patients) using the surface only and this initial transformation was almost identical to registration using the identity matrix as the initial rotation.

Although the improvement in registration accuracy achieved in this paper by combining points and surfaces is statistically significant, the improvement appears to be small (see Fig. 3 and Table I). Part of the observed TRE is due to imperfect target localization, and thus the improvement in true TRE is masked by TLE. Specifically, if TLE is the same in the two image spaces, and is uncorrelated, then

\[(TRE_a)^2 \approx (TRE_s)^2 + 2(TLE)^2\]  

and thus

\[TRE_t = \sqrt{(TRE_a)^2 - 2(TLE)^2}\]  

where TREa and TREs are the observed and true TRE's, respectively. Table II lists true TRE's calculated using this equation with TLE = 0.4 mm (see Section III-C) for registrations using four points only, the surface only, and four points plus the surface. This table illustrates that the fractional improvement in the true TRE (51%) is more impressive than the fractional improvement in the observed TRE (21%).

We demonstrated the efficacy of the WGF algorithm by registering CT and MR volume head images using fiducial points (centroids of rigidly attached external markers) and an anatomical surface. But the potential usefulness of the algorithm is not limited to this specific application. The algorithm can be used with a variety of geometrical information including topological singularities (anatomical landmarks and surfaces) and geometrically invariant differential structures such as crest lines [21] and ridges [55]. Though anatomical landmarks have greater positional uncertainty than our marker centroids, internal landmark registration has the potential to be as accurate as external marker registration if more internal landmarks are used.6 Thus it is possible that registration using several internal landmarks plus a surface will be as accurate as registration using one marker plus a surface, which is more accurate than registration using a surface only. We note that other investigators have performed registration using anatomical landmarks plus a plane [18], [43].

Image-to-physical registration is frequently performed using anatomical surfaces [32], [46]. Intraoperative surface information is sometimes restricted to a small patch due to surgical conditions (e.g., surgical drapes). Surface-based registration using limited surface information is almost certainly not very accurate. The WGF algorithm can be used to improve registration accuracy in such circumstances by incorporating a single point (intrinsic or extrinsic) that effectively serves as an "anchor."

The WGF algorithm might also be useful for extracranial applications, e.g., staging lung cancer with CT, MR, and PET images [58]. Points (e.g., vascular bifurcations), curves (e.g., vascular segments), and surfaces (e.g., pleural surfaces) are plentiful and easily identified. Whereas we used a rigid-body transformation to register head images, it may be more appropriate to use an affine transformation (e.g., rigid-body plus isotropic scaling, rigid-body plus anisotropic scaling, or full affine including shearing effects) to register thoracic or abdominal images because of soft tissue deformations. Registration using a combination of geometrical features may be especially useful in this case since the affine transfor-

6Numerical simulations of point-based registration show that, for a fixed number of points, TRE is roughly proportional to TLE (see footnote 2) [40]. For a fixed TLE, TRE is approximately inversely proportional to the square root of the number of fiducials used, e.g., quadrupling the number of fiducials used will approximately halve the TRE.
mation has more degrees of freedom (parameters) than the rigid-body transformation. \(^7\) We note that the WGF algorithm presented in this paper can be easily extended to use an affine transformation.

Our segmentation method is very flexible and has been used on many parts of the brain, including the skin, brain surface, ventricles, and cerebellum in CT, MR (T1-weighted, PD-weighted, and T2-weighted Spin-Echo and MPRAGE Gradient-Echo), and PET images. We believe that it is also useful for a wide variety of extracranial applications. Specifically, we note that our segmentation method has been used successfully on the heart, spine, bone, and muscle groups in extremity cross sections. These different applications require only changes in the weights of the cost function. The triangulation algorithm we use is rather simple and will work only with simple closed contours that are of similar size and shape in adjacent slices. We are currently implementing a more sophisticated algorithm \(^22\) that should provide adequate results on structures more complicated than the upper part of the head.

This work represents the first determination of TRE for surface-based registration of medical images using Besl and McKay’s ICP algorithm (the WGF algorithm reduces to this case when we use a normalized point set weight of 0.0). The surface-based TRE (mean ± SD = 0.73 ± 0.33 mm) observed in this study is substantially better than that which we previously reported using an independent implementation of the Pelizzari and Chen technique \(^37\), and is substantially better than that reported by a number of other investigators using this latter technique \(^25\), \(^45\), \(^50\), and \(^54\). The fact that we corrected for geometrical distortion in MR images or the fact that we used images that were acquired with a stereotactic frame attached\(^8\) might explain why the accuracy in this paper is better than that reported by other investigators who did not correct for distortion or use images acquired with a frame, but it cannot explain why the accuracy is better than that which we previously reported. It is possible that surface-based registration using the Besl and McKay ICP algorithm is superior to registration using the Pelizzari and Chen technique. It is also quite possible that the superior results in this paper are due to better segmentation. We are currently comparing the accuracy of surface-based registration using various correspondence functions, including the closest point operator, the distance transform approach, and the Pelizzari and Chen technique, on a common set of images. We are also investigating the effect of the accuracy of surface segmentation on the accuracy of registration.

Surface-based registration assumes that the surfaces in the two images are identical or approximately identical. In CT, we find contours at a bright-dark (outer–inner) interface that corresponds physically to the inner surface of the skull. In MR, we find contours at a bright-dark (inner–outer) interface that corresponds physically to a “brain”–bone interface. Mov-

\(^7\) The full affine transformation has 12 parameters (three for the translation vector, nine for the affine matrix) whereas the rigid-body transformation has six (three for the translation vector, three for the rotation matrix).

\(^8\) The frame is an effective head fixation device and may have helped provide better images by immobilizing the patient’s head.

\(^1\) IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 15, NO. 6, DECEMBER 1996

APPENDIX A
CLOSEST POINT SEARCH

The most computationally expensive step in the registration process is finding the closest points. Given a data shape point set \(P\) with \(N_p\) points and a model shape \(X\) with \(N_x\) geometrical primitives (points, line segments, and triangles), the computational complexity of finding the closest points using an exhaustive search is \(O(N_p N_x)\). We use a multidimensional binary search tree \((k-d\,tree, \text{where } k \text{ is the dimensionality of the search space}) to speed up the search process \([4]\) and \([49]\). In our case, \(k = 3\).

The \(k-d\,tree\ is a binary tree in which each node represents a subset of the data records (in our case, points) in a file (in our case, point set) and a partitioning of that subset. Each nonterminal node has two children that represent the two subsets defined by the partitioning. The terminal nodes represent mutually exclusive small subsets of the records (points) called
buckets. We use a bucket size of one. A $k$-$d$ tree divides space into a collection of rectangular parallelepipeds that correspond to the terminal nodes. This data structure provides an efficient method for examining only those points closest to a given point. A $k$-$d$ tree can be constructed in $O(N_d \log N_d)$ time. Each closest point search can be performed in $O(\log N_d)$ time. Thus the computational complexity of finding $N_p$ closest points using a $k$-$d$ tree is $O(N_p \log N_d)$. We construct and search the $k$-$d$ tree using algorithms similar to those presented in [16].

The $k$-$d$ tree can be used with point sets but not with set triangles. In order to find the closest point on a triangle set, we construct a $k$-$d$ tree from the vertices. A search of this tree returns the closest vertex. We assume that the closest surface point lies within one of the triangles of which the vertex is a member. Thus we examine each of these triangles. We have found that this approach produces the actual closest point on the surface, as determined by an exhaustive search, approximately 99.5% of the time. The occasional errors are small and do not affect the performance or accuracy of the registration. We note that the proof of the WGF registration algorithm convergence theorem uses the fact that the closest point operator by definition reduces or leaves unchanged the distance for each point in each iteration [see (8)]. In order to guarantee convergence, we verify that (8) is satisfied for each point in each iteration. In the rare case that (8) is not satisfied, we perform an exhaustive search to find the actual closest point. Alternatively, we could use the closest point from the previous iteration.

**APPENDIX B**

**WEIGHTED POINT-BASED REGISTRATION**

In this appendix, we show that the weighted point-based registration algorithm in Section II-F produces the rigid-body transformation $T$ that minimizes the cost function in (10).

First, we rewrite the expression under the square root sign in (10), using the explicit form of the rigid-body transformation $T(p_j) = R p_j + t$, as

$$
\sum_{j=1}^{N} w_j ||x_j - R p_j - t||^2 .
$$

(24)

Expanding this expression using the definitions for $\bar{p}$ and $\bar{x}$ given in (12), we obtain

$$
\sum_{j=1}^{N} w_j ||x_j - R p_j||^2 + W ||t||^2 - 2Wt \cdot (\bar{x} - \bar{R}p) .
$$

(25)

where $W = \sum_{j=1}^{N} w_j$. By “completing the square” with regard to the second and third terms we obtain

$$
\sum_{j=1}^{N} w_j ||x_j - R p_j||^2 - W ||x - \bar{R}p||^2 \\
+ W ||t - (\bar{x} - \bar{R}p)||^2 .
$$

(26)

The formula for $t$ in (16) minimizes the last term to zero. Since we wish to minimize (26), and since $t$ occurs only in the last term, (16) gives the optimal $t$. By setting the last term to zero, and by using the fact that the mean of $R p_j$ is equal to $R \bar{p}$, we convert (26) to

$$
\sum_{j=1}^{N} w_j ||x_j - R p_j - (\bar{x} - \bar{R}p)||^2 .
$$

(27)

This is easily seen by noting that (26), after setting the last term to zero and dividing through by $W$, has the form

$$
| ||x_j - R p_j||^2 - || \text{mean}(x_j - R p_j)||^2 |
$$

and is thus equal to the variance of $(x_j - R p_j)$. Equation (27) is expanded, using the definitions of $p'_j$ and $x'_j$ given in (11), and using the fact that $||R p'_j|| = ||p'_j||$ for any rotation matrix $R$, to obtain

$$
\sum_{j=1}^{N} w_j ||x'_j||^2 + \sum_{j=1}^{N} w_j ||p'_j||^2 - 2 \sum_{j=1}^{N} w_j (R p'_j) \cdot x'_j .
$$

(29)

Since $R$ occurs only in the last term, minimizing this expression is equivalent to maximizing

$$
\sum_{j=1}^{N} w_j (R p'_j) \cdot x'_j .
$$

(30)

Arun et al. [2] showed that the rotation matrix calculated in Steps 2)-4) of the algorithm in Section II-F maximizes (30) for the unweighted case, i.e., $w_j = 1$. Their proof can be trivially extended to the weighted case by replacing each $p'_j$ by $w_j p'_j$. This extension, plus the derivation above of the optimal $t$, completes our proof of the weighted point-based registration algorithm in Section II-F.

**ACKNOWLEDGMENT**

The authors wish to thank M. Y. Wang and Y. Ge, Ph.D. for their generous assistance and for their many enlightening comments and ideas.

**REFERENCES**


