Comparison and Evaluation of Retrospective Intermodality Brain Image Registration Techniques

Jay West, J. Michael Fitzpatrick, Matthew Y. Wang, Benoit M. Dawant, Calvin R. Maurer, Jr., Robert M. Kessler, Robert J. Maciunas, Christian Barillot, Didier Lemoine, André Collignon, Frederik Maes, Paul Suetens, Dirk Vandermeulen, Petra A. van den Elsen, Sandy Napel, Thilaka S. Sumanaweera, Beth Harkness, Paul F. Hemler, Derek L. G. Hill, David J. Hawkes, Colin Studholme, J. B. Antoine Maintz, Max A. Viergever, Gregoire Malandain, Xavier Pennec, Marilyn E. Noz, Gerald Q. Maguire, Jr., Michael Pollack, Charles A. Pelizzari, Richard A. Robb, Dennis Hanson, Roger P. Woods

December 9, 1996

Short title: Comparison of Brain Image Registration Techniques.

J. West, J. M. Fitzpatrick, and M. Y. Wang are with the Dept. of Computer Science; B. M. Dawant is with the Dept. of Electrical Engineering; C. R. Maurer, Jr. is with the Dept. of Biomedical Engineering; R. M. Kessler is with the Dept. of Radiology; R. J. Maciunas is with the Dept. of Neurological Surgery, Vanderbilt University, Nashville, TN.

C. Barillot and D. Lemoine are with Université de Rennes, Rennes, France; A. Collignon, F. Maes, P. Suetens, and D. Vandermeulen are with Katholieke Universiteit, Leuven, Belgium; P. A. van den Elsen, S. Napel, and T. S. Sumanaweera are with Stanford University Medical Center, Stanford, CA; B. Harkness and P. F. Hemler are with Bowman Gray School of Medicine, Winston-Salem, NC; D. L. G. Hill, D. J. Hawkes, and C. Studholme are with UMDS, Guy's & St. Thomas' Hospitals, London, U.K.; J. B. A. Maintz and M. A. Viergever are with University Hospital Utrecht, Utrecht, The Netherlands; G. Malandain and X. Pennec are with INRIA, Sophia Antipolis, France; M. E. Noz, G. Q. Maguire, Jr., and M. Pollack are with New York University Medical Center, New York, NY; C. A. Pelizzari is with University of Chicago Hospital, Chicago, IL; R. A. Robb and D. Hanson are with Mayo Clinic, Rochester, MN; and R. P. Woods is with UCLA School of Medicine, Los Angeles, CA.

Send all correspondence to: J. M. Fitzpatrick, Department of Computer Science, Vanderbilt University, Box 1679 Station B, Nashville, TN 37235, U.S.A. Phone: 615-322-2365. FAX: 615-343-5459. Email: jmf@vuse.vanderbilt.edu.

Abstract

Objective: The primary objective of this study is to perform a blinded evaluation of a group of retrospective image registration techniques using as a gold standard a prospective, marker-based registration method. In order to ensure blindedness, all retrospective registrations were performed by participants who had no knowledge of the gold-standard results until after their results had been submitted. A secondary goal of the project is to evaluate the importance of correcting geometrical distortion in MR images by comparing the retrospective registration error in the rectified images, i.e., those which have had the distortion correction applied, with that of the same images before rectification.

Materials and Methods: Image volumes of three modalities—CT, MR, and PET—were obtained from patients undergoing neurosurgery at Vanderbilt University Medical Center on whom bone-implanted fiducial markers were mounted. These volumes had all traces of the markers removed and were provided via the Internet to project collaborators outside Vanderbilt, who then performed retrospective registrations on the volumes, calculating transformations from CT to MR and/or from PET to MR. These investigators communicated their transformations again via the Internet to Vanderbilt, where the accuracy of each registration was evaluated. In this evaluation, the accuracy is measured at multiple "volumes of interest" (VOIs), i.e., areas in the brain that would commonly be areas of neurological interest. A VOI is defined in the MR image and its centroid $\mathbf{c}$ is determined. Then, the prospective registration is used to obtain the corresponding point $\mathbf{c}'$ in CT or PET. To this point, the retrospective registration is then applied, producing $\mathbf{c}''$ in MR. Statistics are gathered on the *target registration error* (TRE), which is the distance between the original point $\mathbf{c}$ and its corresponding point $\mathbf{c}''$.

Results: This paper presents statistics on the TRE calculated for each registration technique in this study and provides a brief description of each technique and an estimate of both preparation and execution time needed to perform the registration.

Conclusion: Our results indicate that retrospective techniques have the potential to produce satisfactory results much of the time but that visual inspection is necessary to guard against large errors.

Index Terms: Medical imaging, image registration, fiducial markers, accuracy validation, MR distortion, CT, MR, PET.

I. Introduction

All retrospective image registration methods have attached to them some intrinsic estimate of registration error. However, this estimate of accuracy may not always give a true representation of the distance between actual and estimated positions of targets within the cranial cavity, i.e., the *target registration error* (TRE). By using a method based on fiducial markers [1], [2], [3] as a "gold standard," and by rendering the fiducial markers invisible, we are able to perform a detailed, blinded evaluation of retrospective techniques based on their TRE at several landmark locations within the brain.

Two registration tasks were evaluated in this study, CT to MR and PET to MR, and these tasks were broken into subtasks according to the type of MR and to whether or not the MR image was corrected for geometrical distortion. The image data sets of nine patients were used, seven of which contained both CT and MR scans, and seven with both PET and MR. Spin-echo MR scans of three types were included—T1-weighted (T1), proton-density-weighted (PD), and T2-weighted (T2). Before imaging, each patient had four markers implanted and a COMPASS stereotactic frame attached. For some of the patients, scans were also used that had been corrected for geometrical distortion [4], [5]. The first step toward evaluation is the calculation of the "answers": registrations derived with the aid of the fiducial markers. Next, it is necessary to process the images by removing all traces of the markers and the frame in order to ensure that all subsequent registrations are truly retrospective in nature. We call this process "air brushing."

These air-brushed image volumes were then made available from a central site via the Internet file transfer protocol (FTP) so that each of twelve other sites could apply its own retrospective technique or techniques. The resulting transformations were reported to the central site by electronic mail. A measurement of error was made by calculating the error relative to the gold standard over a set of specified regions known as *volumes of interest* (VOIs).

II. Materials and Methods

A. Image Acquisition

All images were stored in 16 bit, two's complement format:

- The CT volumes have a resolution in the x and y directions of 512 pixels, and between 28 and 34 slices in the z direction. The voxel size is 0.65 mm in x and y, and 4.0 mm in z.
- The MR volumes have a resolution of 256 pixels in the x and y directions, and 20 to 26 slices. The voxel size was between 1.25 and 1.28 mm in the x and y directions, and 4.0 mm in z.
- The PET volumes have a resolution of 128 pixels in the x and y directions, and 15 slices. The voxel size is 2.59 mm in x and y, and 8.0 mm in z.

The CT images were acquired using a Siemens DR-H scanner, the MR images using a Siemens SP 1.5 T scanner, and the PET images with a Siemens/CTI ECAT 933/08–16 scanner. For MR, T1, PD, and T2 images were acquired. The T1 image volumes were acquired with an echo time (TE) of 15 ms and a repetition time (TR) of 650 ms (20 slices) or 800 ms (26 slices), the PD with TE of 20 ms and TR of 2550 ms (20 slices) or 3000 ms

(26 slices), and the T2 with TE of 90 ms and the same TR as PD. Readout gradient strength for T1 was 2.45 mT/m and for PD/T2 was 1.47 mT/m, with four acquisitions in T1 and two in PD/T2. All MR images used half-Fourier reconstruction and a slice selection gradient of 6.8 mT/m. Three additional MR images were acquired for each patient with the identical imaging parameters, except that the readout gradient was reversed. For PET, each patient was injected with 10 mCi of ^{18}F-fluorodeoxyglucose. Scanning was started 40–50 minutes after injection and continued for 25 minutes. Image reconstruction was performed using a Hamming reconstruction filter, resulting in images with a full width at half-maximum resolution of 9 mm.

Internal Review Board authorization was obtained for the use of the patient data sets in this study. All patients whose images were to be used signed a release form indicating their consent.

B. Geometrical Distortion Correction

We correct MR images for static field inhomogeneity by using the image rectification technique of Chang and Fitzpatrick [4], [5]. A new image, without inhomogeneity distortion, is generated from the pair of distorted images acquired with reversed readout gradients. The MR images are corrected for scale distortion by using the COMPASS stereotactic frame as an object of known shape and size. The scale factors are handled by changing the voxel dimensions in the image header.

C. Fiducial Markers and Fiducial Localization

Markers are filled with an aqueous solution designed to be bright in CT and MR. They are constructed from hollow plastic cylinders with an inside diameter of 7 mm and an inside height of 5 mm and attached to plastic marker bases that are screwed into the outer table of the skull of the patient [5], [6]. Markers of an identical construction are used for PET scans but are filled with an aqueous solution of a positron emitter to allow them to be seen clearly in the PET image.

We calculate an intensity-based centroid for each marker using the localization technique described in Wang *et al.* [6]: we call the determination of this position *fiducial localization*. This iterative, knowledge-based technique automatically finds the lowest threshold such that an object formed from voxels whose intensities are higher than the threshold and that are three-dimensionally connected to a selected voxel is neither too large nor too small to be a marker. The large point spread function increases the effective size of the marker substantially in PET. The published localization algorithm therefore was modified for this project to take advantage of the effective size in determining the centroid in the direction perpendicular to a slice.

D. Fiducial-based Registration

When we use markers to register images, we call them *fiducial markers* and call their positions *fiducial points* or *fiducials*. We register CT to MR, and PET to MR, by calculating the rotation and translation parameters of the rigid-body transformation that minimizes the mean square distance between corresponding fiducials in the two images [1], [2]. We have implemented the closed-form solution developed by Arun, Huang, and

Blostein [7]. We define the *fiducial registration error* (FRE) as the root-mean-square (rms) distance between corresponding fiducials after registration and transformation. Marker-based registration was carried out for each patient data set and each MR modality. The derived transformations were tabulated and used as a gold standard for evaluation of the retrospective registration methods.

E. Removal of Fiducial Markers and Stereotactic Frame

The next step is the removal of all traces of the fiducial markers and the stereotactic frame from each image. We call this process "air brushing." This was achieved by manual outlining of regions containing these structures, followed by an approximate reconstruction of the image background in each missing region R. In MR, where the background consists of unstructured noise, pixels at random positions between the edge of R and the lateral image boundary were sampled and placed in R. In CT, as the outer regions of the image are comprised mainly of reconstruction artifacts which take the form of quasi-radial stripes, the approach taken was to interpolate these stripes within the removed region. This was done at a given point P in region R by the following method (Fig. 2):

1. Calculate the radial distances e and i of the point P, respectively, from the external and internal boundaries of the region.
2. Identify points E and I at radial distances e and i, respectively, external and internal to the region. If E lies outside the image, set E to be on the border of the image. A similar precaution is taken to insure that I does not lie within the head.
3. Let the intensity at E be I_e and at I be I_i. Assign P's intensity I_p as $I_p = I_e$ with probability ρ, or $I_p = I_i$ with probability $1 - \rho$, where $\rho = i/(i + e)$.

A similar technique was used for the PET images; in this case, however, pixels in region R were set to an intensity value linearly interpolated between the intensity of the internal and external boundaries of R.

This process is illustrated in Figure 3. The top row (Fig. 3*a*–*c*) shows slices of original volumes from each of the three modalities. The window and level have been set so that the background artifacts may be seen. The middle row (Fig. 3*d*–*f*) shows the same slices after the region R has been outlined and zeroed. This procedure is applied to each slice in the volume. The bottom row (Fig. 3*g*–*i*) shows the slices after reconstruction of the background in the region R. For the MR case, it can be seen that the replaced area is indistinguishable from the rest of the background. In CT and PET, there are slight discontinuities in the direction of the stripes, but the intensity changes smoothly.

For each air-brushed image volume, an ASCII header for each volume was prepared, giving information about the corresponding image, *e.g.* pixel size, resolution, data format and acquisition information. The headers conform to the Interfile standard [8], [9].

F. Communications

After creation of the images and headers was completed, a login name, password, and Internet address were provided to all project participants by e-mail, so that they were able to transfer the images and headers to their own sites by means of FTP.

In order to communicate the retrospective registrations to Vanderbilt once they had been completed, the following scheme was adopted. In the "From" volume (e.g., CT in the case of CT-to-MR registration), the positions of eight points are calculated; taking the origin to be the center of the first voxel in the volume (i.e., the top left pixel of slice zero), the x, y, and z coordinates of the centers of the eight corner voxels in the volume were derived (Fig. 4). These positions were provided via FTP by Vanderbilt for every CT and PET volume in the form of a partially completed "transformation table" for each pair of volumes.

After the retrospective registration transformation was determined, the transformed positions of these eight points *relative to the origin of the "To" modality* (Fig. 5) were computed by each site and entered into the transformation table (as depicted in Figs. 4 and 5, the field of view of the two volumes is typically different, so it is important to specify the volume that provides the origin relative to which the transformed positions are calculated.)

All coordinates were specified to at least four decimal places in units of millimeters. Such high precision insures that any round-off error inherent in converting between a registration transformation and the eight-point sets is negligible. In order to convert the transformation table to a rigid-body transformation, the two point sets are registered using the point-based registration algorithm described in Section II–D. Only three points are necessary to uniquely specify such a transformation, but the full set of eight was used for reasons of symmetry, error reduction, and error prevention.

Clearly, this method of data transmission allows only rigid-body transformations to be accurately communicated, since any nonrigid transformation would be approximated by a rigid one. The use of this protocol thus limits the scope of this project to an evaluation restricted to rigid-body transformations. However, by measuring the FRE (that is produced by the point-based registration algorithm), it is possible to determine whether or not the retrospective registration uses nonrigid deformations, since the FRE would be on the order of millimeters, as opposed to hundredths of a millimeter, for a well-specified rigid-body transformation. This feature is facilitated by the use of a larger point set than necessary, and guards against the possibility of a nonrigid transformation being mistakenly supplied and evaluated as if it were rigid.

Each transformation was transmitted to Vanderbilt by e-mailing an ASCII file containing a completed transformation table.

G. The Retrospective Techniques

The retrospective registrations were performed in parallel at several sites outside Vanderbilt. Some methods were used that were applicable only to CT-to-MR or PET-to-MR registration, and some were suitable for both cases.

1. Barillot and Lemoine [10] used a two-stage technique both for CT-to-MR and PET-to-MR registration. The first step is to perform an approximate registration of objects which are similar in each image. The second stage is the application of a multiresolution Powell [11] algorithm which minimizes the Euclidean distance between the two surfaces given by a $5 \times 5 \times 5$ chamfer mask. The total time taken for each registration

was approximately 15 minutes using a Sun SPARCstation 20.

2. Collignon, Maes, Suetens, and Vandermeulen [12] used a fully automatic technique for CT-to-MR and PET-to-MR registration. The technique employs a Powell algorithm to search the space of registration parameters, judging the images to be in perfect registration when the mutual information of the intensities of corresponding voxels is maximized. Each registration took 10 to 30 minutes in the CT-to-MR case, and 2 to 4 minutes in the PET-to-MR case, using an IBM RS6000/3AT.
3. Van den Elsen, Napel, and Sumanaweera [13] used a fully automated multiresolution gray-value correlation technique for registering CT volumes to MR. The total time taken for each registration was approximately 1 to 2 hours using an SGI Indigo 2.
4. Harkness applied a surface-matching technique developed by Pelizzari *et al.* [14] to CT-to-MR and PET-to-MR registration. The typical time required for each registration was approximately 20 minutes, most of which was human interaction time.
5. Hemler, Napel, and Sumanaweera [15] employed a surface-based method for registering from CT to MR. In this system, the corresponding surface to be registered is first identified in each image set. This process results in a triangular mesh representing the surface in one image set. and a set of 2-D polygon points are used to represent the surface in the other image set. A least-squares minimization technique is then used to determine the rigid-body transformation which minimizes a cost function related to the sum-square perpendicular distance between the two surfaces. (No estimate of registration time was provided by these investigators.)
6. Hill, Hawkes, and Studholme [16], [17], [18] registered CT to MR and PET to MR using a fully automated multi-resolution voxel similarity measure algorithm based on the mutual information of the joint probability distribution (proposed by Collignon [12] and Wells [19].) The time taken for each registration was approximately 20 minutes on a Sun SPARCstation 20/61.
7. Maintz, van den Elsen, and Viergever [20], [21], [22] registered CT to MR and PET to MR by matching "edgeness" volumes. The method is fully automated. The algorithm used to match the edgeness volumes is described in [20], [21], [22]. All computations were done on HP 700-series workstations. For each registration, the feature extraction typically takes a few minutes. The matching itself can take from 1 to 12 hours for CT-to-MR registration and about 40 minutes for PET-to-MR registration, but these times may be decreased by using larger step sizes and omitting fail-safe measures.
8. Malandain and Pennec [23], [24], [25] registered CT to MR and PET to MR. The method was originally designed for PET-to-MR registration and is used for CT-to-MR registration without modification. The first step is the segmentation of the brain in both images by means of morphological and topological operators. This is done automatically, with human inspection and possible changes in the parameters if a satisfactory result is not achieved. The second step is the construction of a potential energy field of one volume, and the computation of the global force and torque acting on the other volume due to this field; these values are used to bring the two volumes into alignment. Each registration took approximately 3 minutes for CT-to-MR registration and 40 seconds for PET-to-MR registration using a DEC Alpha 3000.

9. Noz, Maguire, and Pollack [26] used a two-step method for registration, both for CT to MR and PET to MR. First, an oblique projection, which involves rotating the 3-D data set in the coronal and sagittal planes by angles determined by visual inspection, is performed in order to bring the image slice planes into coincidence. Eight to 12 landmarks pairs are then chosen on a corresponding slice from each image volume, and a first- or second-order polynomial (warping) transformation is applied whose coefficients minimize the rms distance between the point sets. The time typically required for each registration was between 15 and 30 minutes using a Sun SPARCstation 10/51, most of which was spent finding corresponding landmarks.
10. Pelizzari [14] used a two-step method for CT-to-MR and PET-to-MR registration. The first stage was the segmentation of contours using a combination of manual and threshold-based boundary following. The contours in the "From" modality were represented as a set of points and in the "To" modality as a stack of disks. The second stage was the minimization, by means of a Powell algorithm, of the mean square distance of the points from the surface, measured by finding the intersection of the surface model with a ray drawn from each point to the centroid of the surface model. The typical time taken for each registration was 20 minutes; almost all of this spent in the segmentation step. The CPU time required for registration was approximately 5 seconds on an SGI Indigo 2.
11. Robb and Hanson [27], [28] used a multiresolution, multithreshold, surface matching algorithm based on parametric chamfer matching for CT-to-MR and PET-to-MR registration. Four variations were performed using two different surfaces and two numbers of points per surface (RO1/RO2 were performed by matching skin to skin, RO3/RO4 by matching brain—or inner surface of skull in CT—to brain; RO1/RO3 were performed using 100 points, RO2/RO4 were performed using 1,000 points). Surface segmentation was accomplished using an automatic 3-D mathematical morphological algorithm. The total time for each registration on an SGI Challenge workstation ranged from 3 to 20 minutes, with morphologic-based segmentation taking 2 to 3 minutes, manual editing in selected cases taking 10 to 15 minutes, and actual registration taking 30 seconds (100 points) to 2 minutes (1,000 points).
12. Woods [29] registered PET to MR using a multi-sampling density method based on the correlation of intensity values between PET and MR. Two methods were used. In the first method (WO1), nonbrain regions (i.e., skin, skull, meninges) are manually removed. After this manual editing, automated registration is performed by dividing the MR image into 256 partitions based on intensity. A Newton-Raphson method is used to find the transformation parameters that minimize a weighted average of the standard deviations of PET voxel intensities corresponding to each MR partition. The second method (WO2) is identical to the first except that the MR images are edited more extensively to remove areas of focal brain pathology. The total time for each registration was typically between 20 and 30 minutes using a Sun SPARCstation IPX, with manual editing taking 15 to 20 minutes and cost function minimization taking less than 5 minutes.

H. Data Analysis

Each submission of retrospective transformations was accompanied by a statement indicating the cases in which the registration procedure was felt to have failed, i.e., was not good enough to be clinically useful. In all cases, the statements indicated that the registration was successful on every data set provided.

At Vanderbilt, after the transformation tables have been received from each site and the corresponding rigid-body transformations have been determined, the next step in the evaluation is to perform a comparison between these registrations and the fiducial-based ones. In collaboration with a neurological and a neurosurgical expert, a set of VOIs representing areas of neurological and/or surgical interest was manually segmented within one of the MR image volumes for each patient. The VOIs were stored as sets of x, y, and z voxel coordinates.

An estimate of the accuracy of the retrospective registration at the position of each VOI is computed as follows (Fig. 1). The centroid pixel of the VOI is found, and its position is converted from a voxel index to a millimetric position $\mathbf{c}$ in the "To" modality using the known voxel size for the image volume. Let $\mathbf{R}_g$ and $\mathbf{t}_g$ be the rotation matrix and translation vector, respectively, of the gold-standard rigid-body transformation G, and $\mathbf{R}_r$ and $\mathbf{t}_r$ be the rotation and translation components of the retrospective transformation R. The point $\mathbf{c}'$ in the "From" modality is defined so that $\mathbf{c}$ is the mapping of $\mathbf{c}'$ under the gold-standard transformation. Thus,

$$\mathbf{c} = G(\mathbf{c}') = \mathbf{R}_g\mathbf{c}' + \mathbf{t}_g. \tag{1}$$

By inverting equation (1), we obtain

$$\mathbf{c}' = G^{-1}(\mathbf{c}) = \mathbf{R}_g^{-1}\mathbf{c} - \mathbf{R}_g^{-1}\mathbf{t}_g. \tag{2}$$

The point $\mathbf{c}''$ in the "To" modality is defined as the mapping of $\mathbf{c}'$ under the retrospective transformation. Thus,

$$\mathbf{c}'' = R(\mathbf{c}') = \mathbf{R}_r\mathbf{c}' + \mathbf{t}_r. \tag{3}$$

The difference between the registered target position of the retrospective method and that of the gold standard is $\mathbf{d} = \mathbf{c}'' - \mathbf{c}$. We define the TRE of the retrospective registration at the anatomical location of the VOI to be the Euclidean distance d between $\mathbf{c}$ and $\mathbf{c}''$, i.e., $\mathrm{TRE} = d = ||\mathbf{d}||$.

In Appendix A, we examine the spatial dependence of TRE. If the retrospective transformation differs from the gold standard by only a translation, then it is easy to see that $\mathrm{TRE} = ||\mathbf{t}_r - \mathbf{t}_g||$ everywhere. For the more common case, in which the difference includes a rotational component, the spatial dependence of TRE is only slightly more complicated. We show that TRE exhibits a simple cylindrical symmetry about an axis $\mathbf{s}$ running through the "To" space as follows:

$$\mathrm{TRE} = \sqrt{ar^2 + b^2}, \tag{4}$$

where $a > 0$ and b are parameters that are independent of position, and r is the perpendicular distance of $\mathbf{c}$ from $\mathbf{s}$.

Using the formulas in Appendix A, we could calculate a, b, and $\mathbf{s}$ for a given retrospective transformation and gold-standard transformation. Large values for a and b and/or a large distance between $\mathbf{s}$ and the center of the image volume would then imply large values for TRE. It is difficult, however, to translate these values into a clinically meaningful quantity. Instead, we have chosen to measure TRE directly at a discrete set of centroids of anatomically meaningful VOIs. While we do not use it in our measurements of TRE, equation (4) still provides an important underpinning for our method of evaluation. The smoothness of the simple r^2 variation in this equation assures us that our discrete sampling will not cause us to miss large values of TRE.

The particular anatomical positions corresponding to the VOIs used in this evaluation are as follows: 1) maximum aperture of fourth ventricle, 2) junction of fourth ventricle with aqueduct, 3) right globe, 4) left globe, 5) optic chiasm, 6) apex of left Sylvian fissure, 7) apex of right Sylvian fissure, 8) junction of central sulcus with midline, 9) left occipital horn, and 10) right occipital horn.

III. Results

For a given technique, the TREs are very similar in all of the VOIs indicating that the r^2 dependence of equation (4) is masked by variations in a, b, and the position and orientation of the rotational axis $\mathbf{s}$. The similarities are such that we are reporting statistics only for the pooled VOIs. A complete listing of the results is available via FTP (see the project web homepage[1]).

Tables 1 through 4 present the median and maximum TREs for the CT-to-MR and PET-to-MR registrations. We have chosen to present medians instead of means to reduce the influence of outliers. The median and maximum TREs listed for each technique and registration modality pair are the median and maximum of all individual TREs for all applicable patients and VOIs for that technique and modality pair. Each row represents a registration modality pair, and each column represents a registration technique. The sites are represented by abbreviation of the participants' names as follows: 1) Barillot *et al.* (BA), 2) Collignon *et al.* (CO), 3) van den Elsen *et al.* (EL), 4) Harkness (HA), 5) Hemler *et al.* (HE), 6) Hill *et al.* (HI), 7) Maintz *et al.* (MAI), 8) Malandain *et al.* (MAL), 9) Noz *et al.* (NO), 10) Pelizzari (PE), 11–14) Robb *et al.* methods 1–4 (RO1, RO2, RO3, RO4), and 15–16) Woods methods 1 and 2 (WO1, WO2).

We investigated the effects of geometrical distortion correction in MR on registration accuracy by comparing the CT-to-MR and PET-to-MR registration errors with and without correcting for distortion. First, we pooled the results of all registration techniques and used the median TRE values listed in Tables 1 and 3. In all cases the differences are not significant (two-tailed paired t-test, $P < 0.05$). Second, we examined each technique individually and used as data for comparison the median TRE values obtained for each patient with that technique. These differences are significant (two-tailed paired t-test, $P < 0.05$) in only one case: CT-to-T2 registration by Collignon. The differences are "marginally" significant ($P < 0.10$) in four other cases: CT-to-T2 registration by Harkness, and CT-to-T1 registration by Collignon, by van den Elsen, and by Hill.

[1] `http://cswww.vuse.vanderbilt.edu/~image/registration/`

IV. Discussion

The principal goal of this project was to determine the accuracy of retrospective image registration techniques. It should be noted that because this study assesses only image-to-image registration, and not image-to-physical space registration, its direct clinical application lies in intermodality image correlation. Clinical applications might include, for example, the assignment of anatomic specificity to functional activation studies with functional MR imaging (fMRI) and PET or the longitudinal cross-correlation over time of imaging studies to follow tumor growth and response to therapy. In using our results to guide such applications it is important to consider the validity of our approach and the accuracy of our gold standard.

A. Submission Errors

After all transformation tables had been submitted, the gold-standard tables and the original image volumes with visible fiducials were made available so that the correctness of the evaluation could be verified by all participants. In addition, each site was provided with error statistics for its technique(s). At this point, the blinded study was over. After this point, four errors were discovered:

- Hill *et al.* reported that a systematic error was made while converting their transformation format to the Vanderbilt format and submitted revised tables. Specifically, Hill stated that the angle of rotation about the x axis was accidentally negated. Vanderbilt independently verified this error, and the error was corrected (noted by the double dagger in Tables 1 through 4). Because the correction is independently verifiable, the blindedness of the study is not compromised by the inclusion of the corrected data.
- Barillot *et al.* reported that for one patient the wrong set of tables was submitted for the rectified CT-to-MR registrations. In this case, resubmission would not be appropriate. Instead, the three erroneous transformations were dropped from the study (noted by the asterisks in Tables 1 through 2).
- Collignon *et al.* reported that a manual entry error was made while creating one transformation table for one patient. Again, resubmission was not appropriate, and the single erroneous transformation was dropped from the study (noted by the asterisks in Tables 3 through 4).
- Robb *et al.* reported that for one patient the wrong set of tables was submitted for the rectified CT-to-PD skull-to-brain and PET-to-PD brain-to-brain registrations. In this case, resubmission would not be appropriate. Instead, the four erroneous transformations were dropped from the study (noted by the asterisks in Tables 1 through 4).

We include in Appendix B all results that were obtained before correcting (Hill) or dropping (Barillot, Collignon, and Robb) erroneous transformations.

Finally, we note that Maintz *et al.* withdrew from the study after submitting a partial set of results (CT-to-T1 and CT-to-PD unrectified and rectified). After publication of the results for the remaining sites [30], this site was permitted to rejoin the study and to complete the submission. Because the gold-standard transformations and the fiducial-containing images were withheld from this site until after the complete submission, the blindedness of the study is not compromised.

B. Gold-Standard Accuracy

The TREs in this study are exact measures of error only if the gold standard provides perfect registration. If the transformation errors of the standard and a retrospective technique are independent, the following simple relationship holds for the rms of the observed TRE, the true TRE ($\mathrm{TRE_t}$), and the gold-standard TRE ($\mathrm{TRE_g}$):

$$\mathrm{rms[TRE]} = \sqrt{\mathrm{rms^2[TRE_t]} + \mathrm{rms^2[TRE_g]}}. \tag{5}$$

While we cannot measure $\mathrm{rms[TRE_g]}$ directly, we can estimate it using numerical simulations [1], [2], [3]. The simulations show that for four fiducial markers $\mathrm{rms[TRE_g]} \approx 0.94\ \mathrm{rms[FRE]}$. The rms FRE for marker-based CT-to-MR registration using the three types of MR images corrected for geometrical distortion is 0.41 mm. The rms FRE for PET-to-MR registration is 1.75 mm. Thus estimated $\mathrm{rms[TRE_g]}$ is approximately 0.39 mm for CT-to-MR and 1.65 mm for PET-to-MR. The larger TREs for registrations involving PET are to be expected because of the larger voxels in this modality.

These simulations do not account for geometrical distortion and therefore apply only to the registrations obtained using MR images corrected for geometrical distortion. The gold-standard TREs can be expected to be somewhat larger for the uncorrected MR images.

We corrected the MR images for scale distortion by using the stereotactic frame as an object of known shape and size. The frame N-bar cross-sections lie in the image periphery, where gradient nonlinearity is highest [31], but improvements in the FRE obtained using our linear scale corrections suggest that the scale factors used in this study improve the geometrical fidelity in the central portion of the images as well. Except for one patient, the scale factors were greater than 1.004 (i.e., 0.4%).

C. Sensitivity Analysis

A preliminary version of this work was presented at *Medical Imaging 1996* [30]. The current version includes results from an additional site (Maintz *et al.*). The remaining results differ slightly from the preliminary ones. The differences are due to a difference in the set of VOIs used in the current version. The earlier VOIs were lost in a disk failure. The current VOIs were once again manually outlined for each patient. The statistics in Tables 1 through 4 are based on these new VOIs. The mean of the absolute percentage difference in medians between the earlier VOIs and the current ones is 2.9% for CT-MR; for PET-MR, it is 3.5%. These differences are small, but they suggest a sensitivity analysis to determine what aspect of changes in VOIs contributes most to changes in these statistics.

As mentioned in Section II–H, the spatial variation of TRE is smooth, and as mentioned in Section III, the variation among VOIs is small. To measure the sensitivity to VOI position we perturbed the centroid of each VOI randomly. Our experience in outlining the VOIs suggests that the centroid can vary by as much as one voxel. To simulate such variation, we perturbed the current VOIs: in each of 50 runs, the x, y, and z components of the position of each centroid were altered independently with the displacement of each component chosen from a population with zero mean and a standard deviation equal to one voxel's width in the respective direction (i.e., x, y, or z). The mean of the absolute percentage difference in medians as a result of these perturbations is only 1.5% and 1.3% for CT-MR and PET-MR, respectively, indicating that perturbations of centroid positions resulting from imprecise drawing of VOIs will not have an appreciable effect on the evaluation.

There is a second less obvious source of variability from the drawing of VOIs. Because of some variation in the coverage of the head, some VOIs are absent in some patient images and present in others. Furthermore, when drawing VOIs, there is in some cases (near the top and bottom slices) considerable subjectivity in deciding whether a VOI is present or absent. Thus, it is likely that a few VOIs have been either dropped or added in the current work. The small variation among VOIs observed at the beginning of this section suggests that dropping or adding a given VOI will have little consequence. While that may be true when the VOI is dropped or added for all patients, the dropping or adding of a VOI for a given patient has the additional effect of increasing or decreasing the influence of that patient on the median over all patients and all VOIs. Thus, if the results are relatively more sensitive to the presence of a given *patient*, then dropping or adding a VOI or VOIs for that patient may have a relatively larger effect on the median. To investigate the sensitivity to patient selection, we recomputed the medians by dropping patients one by one. We computed the mean absolute percentage change in median over all such patients. The results are 8.9% for CT-MR and 8.1% for PET-MR. These considerably larger changes show that patient selection may be more important than VOI selection and suggest that the observed differences between the earlier version and the present one are caused by the dropping or adding of VOIs in a few patients.

It seems clear from this analysis that the selection of patients and patient images can have an appreciable effect on the mean accuracies achieved in a study of this kind. Thus one should be careful when comparing accuracies based on registration results achieved on different data sets.

D. Observations

There is little variation within a given technique of the median and maximum registration error among the three MR types—T1, PD and T2—used in this study. Also, it is clear that the accuracy of registrations involving CT is better than the accuracy of registrations involving PET, probably because of the larger slice thickness in PET.

As discussed at the end of Section III, there is a significant improvement in registration accuracy after correction of geometrical distortion in MR for only one technique, and a "marginally" significant improvement for four other techniques. We note that a previous study found that the mean point-based TRE using three fiducials improved from 1.2 mm to 0.8 mm and the mean surface-based TRE improved from 2.2 mm to 1.2 mm [5]. It is possible that improvement in accuracy produced by distortion correction is too small to be seen for many of the registrations in this study.

If a retrospective technique's accuracy is approximately the same as the gold standard, then it follows from equation (5) that the rms of the observed TRE will be approximately $\sqrt{2}\ \mathrm{rms}[\mathrm{TRE_g}]$, which is approximately 0.55 mm for CT-to-MR and 2.33 mm for PET-to-MR. We note that the smallest median values are 0.7 mm for CT-to-MR (Table 1) and 2.0 mm for PET-to-MR (Table 3). The smallest rms values (not tabulated) are 1.0 mm for CT-to-MR (Collignon *et al.*, PD rectified; Hill *et al.*, T1 rectified and PD rectified) and 2.3 mm for PET-to-MR (Woods method 1, T1 rectified), suggesting that the accuracy of some of the retrospective techniques approaches the accuracy of the bone-implanted fiducial marker method, at least for PET-to-MR registration. Thus, it appears that retrospective image registration has the potential of providing satisfactory results much of the time.

The TRE maxima are often large, however (see Tables 2 and 4), suggesting that visual inspection should play an important role in the clinical application of registration techniques. If visual inspection reveals a poor registration, then the registration might be repeated (e.g., by using a different initial transformation or by using different registration parameters), or it might not be used at all. In this study, the option for discarding registrations was available but was not exercised by any group. Discarding may have improved the statistics. For example, one group (Malandain *et al.*) reported that their algorithm was run for all pairs of images with the same parameter settings and without visual inspection. Visual inspection by this group or by any group might have revealed poor registrations, which could have been declared to be failures and thereby excluded from the tables.

Though some registration techniques produced smaller errors than others in this study, it is not possible to draw statistically meaningful conclusions regarding the superiority of a technique for a given task (or subtask). That was not a goal of this project and, because of the large number of techniques (16) and the relatively small number of independent observations (maximum of seven patients) used in their evaluation, it is not statistically feasible.

We note finally that Noz *et al.* performed registration using nonrigid transformations. Our evaluation method approximated their transformations with rigid-body transformations. Thus the results reported in this study may not be a fair evaluation of the accuracy of their technique.

V. Conclusion

Twelve groups of investigators applied 16 techniques to selected registration tasks involving the registration of CT and/or PET to MR. Our results indicate that retrospective techniques have the potential to produce satisfactory results much of the time but that visual inspection is necessary to guard against large errors.

VI. Appendices

A. Spatial Variation of TRE

In this appendix, we derive the spatial variation of TRE, equation (4), for the case in which the difference between the retrospective and gold-standard transformations includes a rotational component.

We begin by substituting the expression for $\mathbf{c}'$ derived in equation (2) into equation (3) to get

$$\mathbf{c}'' = R(G^{-1}(\mathbf{c})) = \mathbf{R}_r\mathbf{R}_g^{-1}\mathbf{c} - \mathbf{R}_r\mathbf{R}_g^{-1}\mathbf{t}_g + \mathbf{t}_r. \tag{6}$$

Thus,

$$\mathbf{d} = \mathbf{c}'' - \mathbf{c} = \mathbf{R}_r\mathbf{R}_g^{-1}\mathbf{c} - \mathbf{c} + \mathbf{t}_r - \mathbf{R}_r\mathbf{R}_g^{-1}\mathbf{t}_g \tag{7}$$

or

$$\mathbf{d} = \mathbf{R}\mathbf{c} - \mathbf{c} + \mathbf{t}, \tag{8}$$

where $\mathbf{R} \equiv \mathbf{R}_r\mathbf{R}_g^{-1}$ and $\mathbf{t} \equiv \mathbf{t}_r - \mathbf{R}_r\mathbf{R}_g^{-1}\mathbf{t}_g$. We note that since $\mathbf{R}_g$ is an orthogonal matrix, $\mathbf{R}_g^{-1} = \mathbf{R}_g^t$, where the superscript t denotes transposition.

$\mathbf{R}$ can be completely described in terms of an axis of rotation $\mathbf{v}$ and an angle of rotation θ about $\mathbf{v}$. If we resolve the vectors $\mathbf{c}$ and $\mathbf{t}$ into their components parallel to ($\mathbf{c}_{||}$, $\mathbf{t}_{||}$) and perpendicular to ($\mathbf{c}_\perp$, $\mathbf{t}_\perp$) $\mathbf{v}$, we have

$$\mathbf{d} = \mathbf{R}\mathbf{c}_\perp - \mathbf{c}_\perp + \mathbf{t}_\perp + \mathbf{t}_{||}, \tag{9}$$

where we have used $\mathbf{R}\mathbf{c}_{||} = \mathbf{c}_{||}$. If we now find the particular value of $\mathbf{c}_\perp$, which we will call $\mathbf{c}_\perp^{(0)}$, for which

$$\mathbf{c}_\perp^{(0)} = \mathbf{R}\mathbf{c}_\perp^{(0)} + \mathbf{t}_\perp, \tag{10}$$

and define $\mathbf{r} \equiv \mathbf{c}_\perp - \mathbf{c}_\perp^{(0)}$, we have

$$\mathbf{d} = \mathbf{R}\mathbf{r} - \mathbf{r} + \mathbf{t}_{||}. \tag{11}$$

The meaning of this equation is that each $\mathbf{d}$ is the result of a rotation about an axis $\mathbf{s}$ parallel to $\mathbf{v}$ and passing through $\mathbf{c}_\perp^{(0)}$ followed by a translation parallel to $\mathbf{s}$. Taking the dot product of $\mathbf{d}$ with itself gives

$$d^2 = \mathbf{d}\cdot\mathbf{d} = 2(\mathbf{r}\cdot\mathbf{r} - \mathbf{r}\cdot\mathbf{R}\mathbf{r}) + \mathbf{t}_{||}\cdot\mathbf{t}_{||}, \tag{12}$$

where we have used $\mathbf{r}\cdot\mathbf{t}_{||} = \mathbf{R}\mathbf{r}\cdot\mathbf{t}_{||} = 0$. Noting that TRE $\equiv d$, we have

$$\mathrm{TRE} = \sqrt{ar^2 + b^2}, \tag{4}$$

where $a = 2(1 - \cos\theta)$, $r = ||\mathbf{r}||$, $b = ||\mathbf{t}_{||}||$, θ is the angle of rotation about $\mathbf{s}$, and r is the perpendicular distance of $\mathbf{c}$ from $\mathbf{s}$.

B. Submission Error Results

We list here all results that were obtained before correcting (Hill) or dropping (Barillot, Collignon, and Robb) erroneous transformations (see Section IV–A).

- Hill. Table 1 (column HI, row 1 to row 6): 1.6, 2.0, 1.7, 1.2, 1.2, 1.7; Table 2: 3.4, 4.5, 4.7, 3.0, 3.4, 4.2; Table 3: 2.9, 3.5, 2.2, 3.2, 3.4, 2.4; Table 4: 8.8, 9.8, 8.3, 9.5, 7.4, 8.6.
- Barillot. Table 1 (column BA, row 4 to row 6): 1.8, 2.0, 2.4; Table 2: 22.0, 39.4, 25.8.
- Collignon. Table 3 (column CO, row 4): 3.0; Table 4: 107.0.
- Robb. Table 1 (column RO3 to column RO4, row 5): 5.7, 5.8; Table 2: 173.0, 172.9; Table 3: 4.5, 4.3; Table 4: 176.2, 61.6.

VII. Acknowledgment

A preliminary version of this work was presented at the conference *Medical Imaging 1996* (Newport Beach, CA).

Funding for this project was provided by the National Institutes of Health, Project Number 1 R01 NS33926-02.

The Vanderbilt collaborators on this project would like to thank Yu Shyr, PhD for the time and effort he spent advising us about the statistical methods used in this study; John Votaw, PhD (Emory University, Atlanta, GA) for answering many questions about PET imaging; and Srikanth Gadamsetty for helping to collect and process patient images. They would also like to thank Nicholas Ayache, PhD (INRIA, Sophia Antipolis, France) and Frederic Fahey, DSc (Bowman Gray School of Medicine, Winston-Salem, NC) for their support and encouragement. The investigators from the Mayo Clinic (Rochester, MN) would like to thank Vanessa Murrie for her help in performing registrations. The investigators from New York University Medical Center (New York, NY) would like to thank Elissa Kramer, MD for her help and advice.

References

[1] V. R. Mandava, J. M. Fitzpatrick, C. R. Maurer, Jr., R. J. Maciunas, and G. S. Allen, "Registration of multimodal volume head images via attached markers", *Medical Imaging VI: Image Processing*, vol. Proc. SPIE 1652, pp. 271–282, 1992.

[2] C. R. Maurer, Jr., J. M. Fitzpatrick, R. L. Galloway, Jr., M. Y. Wang, R. J. Maciunas, and G. S. Allen, "The accuracy of image-guided neurosurgery using implantable fiducial markers", in *Computer Assisted Radiology 1995*, H. U. Lemke, K. Inamura, C. C. Jaffe, and M. W. Vannier, Eds., pp. 1197–1202. Springer-Verlag, Berlin, 1995.

[3] C. R. Maurer, Jr., J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, Jr., R. J. Maciunas, and G. S. Allen, "Registration of head volume images using implantable fiducial markers", Tech. Rep. CS96-03, Department of Computer Science, Vanderbilt University, Sept. 1996.

[4] H. Chang and J. M. Fitzpatrick, "A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities", *IEEE Trans. Med. Imaging*, vol. 11, pp. 319–329, 1992.

[5] C. R. Maurer, Jr., G. B. Aboutanos, B. M. Dawant, S. Gadamsetty, R. A. Margolin, R. J. Maciunas, and J. M. Fitzpatrick, "Effect of geometrical distortion correction in MR on image registration accuracy", *J. Comput. Assist. Tomogr.*, vol. 20, pp. 666–679, 1996.

[6] M. Y. Wang, C. R. Maurer, Jr., J. M. Fitzpatrick, and R. J. Maciunas, "An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head", *IEEE Trans. Biomed. Eng.*, vol. 43, pp. 627–637, 1996.

[7] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, pp. 698–700, 1987.

[8] B. S. Baxter, L. E. Hitchner, and G. Q. Maguire, Jr., "A standard format for digital image exchange", *American Association of Physicists in Medicine*, 1982.

[9] T. D. Cradduck, D. L. Bailey, B. F. Hutton, F. de Conninck, E. Busemann-Sokole, H. Bergmann, and U. Noelpp, "A standard protocol for the exchange of nuclear medicine image files", *Nucl. Med. Comm.*, vol. 10, pp. 703–713, 1989.

[10] D. Lemoine, D. Liegeard, E. Lussot, and C. Barillot, "Multimodal registration system for the fusion of MRI, CT, MEG, and 3D or stereotactic angiographic data", *Medical Imaging 1994: Image Capture, Formatting, and Display*, vol. Proc. SPIE 2164, pp. 46–56, 1994.

[11] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives", *Comput. J.*, vol. 7, pp. 155–163, 1964.

[12] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory", in *Information Processing in Medical Imaging 1995*, Y. Bizais, C. Barillot, and R. Di Paola, Eds., pp. 263–274. Kluwer Academic, Dordrecht, The Netherlands, 1995.

[13] P. A. van den Elsen, E.-J. D. Pol, T. S. Sumanaweera, P. F. Hemler, S. Napel, and J. R. Adler, "Grey value correlation techniques used for automatic matching of CT and MR brain and spine images", *Visualization in Biomedical Computing 1994*, vol. Proc. SPIE 2359, pp. 227–237, 1994.

[14] C. A. Pelizzari, G. T. Y. Chen, D. R. Spelbring, R. R. Weichselbaum, and C.-T. Chen, "Accurate three-dimensional registration of CT, PET, and/or MR images of the brain", *J. Comput. Assist. Tomogr.*, vol. 13, pp. 20–26, 1989.

[15] P. F. Hemler, T. S. Sumanaweera, P. A. van den Elsen, S. Napel, and J. R. Adler, "A versatile system for multimodality image fusion", *J. Image Guid. Surg.*, vol. 1, pp. 35–45, 1995.

[16] D. L. G. Hill, D. J. Hawkes, N. Harrison, and C. F. Ruff, "A strategy for automated multimodality registration incorporating anatomical knowledge and imager characteristics", in *Information Processing in Medical Imaging 1993*, H. H. Barrett and A. F. Gmitro, Eds., pp. 182–196. Springer-Verlag, Berlin, 1993.

[17] D. L. G. Hill and D. J. Hawkes, "Voxel similarity measures for automated image registration", *Visualization in Biomedical Computing 1994*, vol. Proc. SPIE 2359, pp. 205–216, 1994.

[18] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3D registration of MR and CT images of the head", *Med. Image Anal.*, vol. 1, 1996 (In press).

[19] W. M. Wells, III, P. Viola, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information", in *Medical Robotics and Computer Assisted Surgery 1995*, pp. 55–62. Wiley-Liss, New York, 1995.

[20] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Comparison of feature-based matching of CT and MR brain images", in *Computer Vision, Virtual Reality, and Robotics in Medicine 1995*, N. Ayache, Ed., pp. 219–228. Springer-Verlag, Berlin, 1995.

[21] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Evaluation of ridge seeking operators for multimodality medical image matching", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, pp. 353–365, 1996.

[22] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Comparison of edge-based and ridge-based registration of CT and MR brain images", *Med. Image Anal.*, 1996 (In press).

[23] G. Malandain, S. Fernández-Vidal, and J. M. Rocchisani, "Improving registration of 3-D medical images using a mechanical based method", in *3rd European Conference on Computer Vision (ECCV '94)*, 1994, pp. 131–136.

[24] G. Malandain, S. Fernández-Vidal, and J. M. Rocchisani, "Rigid registration of 3-D objects by motion analysis", in *Proc. 12th Int. Conf. Pattern Recognition*, 1994, pp. 579–581.

[25] G. Malandain, S. Fernández-Vidal, and J. M. Rocchisani, "Physically based rigid registration of 3-D free-form objects: Application to medical imaging", Tech. Rep. 2453, INRIA, Sophia Antipolis Cedex, France, January 1995.

[26] G. Q. Maguire, Jr., M. E. Noz, H. Rusinek, J. Jaeger, E. L. Kramer, J. J. Sanger, and G. Smith, "Graphics applied to medical image registration", *IEEE Comput. Graph. Appl.*, vol. 11, pp. 20–29, March 1991.

[27] H. Jiang, K. S. Holton, and R. A. Robb, "Image registration of multimodality 3-D medical images by chamfer matching", *Biomedical Image Processing and Three-Dimensional Microscopy 1992*, vol. Proc. SPIE 1660, pp. 356–366, 1992.

[28] H. Jiang, R. A. Robb, and K. S. Holton, "A new approach to 3-D registration of multimodality medical images by surface matching", *Visualization in Biomedical Computing 1992*, vol. Proc. SPIE 1808, pp. 196–213, 1992.

[29] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "MRI-PET registration with automated algorithm", *J. Comput. Assist. Tomogr.*, vol. 17, pp. 536–546, 1993.

[30] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, P. F. Hemler, S. Napel, T. S. Sumanaweera, B. Harkness, D. L. G. Hill, C. Studholme, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, Jr., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, "Comparison and evaluation of retrospective intermodality image registration techniques", *Medical Imaging 1996: Image Processing*, vol. Proc. SPIE 2710, pp. 332–347, 1996.

[31] L. R. Schad, S. Lott, F. Schmitt, V. Sturm, and W. J. Lorenz, "Correction of spatial distortion in MR imaging: A prerequisite for accurate stereotaxy", *J. Comput. Assist. Tomogr.*, vol. 11, pp. 499–505, 1987.

TABLE 1
MEDIAN ERRORS FOR CT-TO-MR REGISTRATION

MR Modality	Technique														N
	BA	CO	EL	HA	HE	HI‡	MAI	MAL	NO†	PE	RO1	RO2	RO3	RO4	
T1	1.6	1.5	1.6	3.4	1.4	1.2	5.1	4.3	3.3	2.7	4.2	5.2	5.7	5.4	7
PD	1.9	1.5	2.0	3.1	2.4	1.9	4.1	4.0	7.8	1.9	4.5	5.5	4.9	4.8	7
T2	2.5	1.5	1.6	4.2	4.7	1.5	3.9	5.0	3.9	2.5	4.5	4.5	5.4	4.7	7
T1 rect.	1.4*	0.7	0.9	3.3	1.0	0.7	4.9	5.4	3.4	2.2	5.9	5.9	6.3	5.9	6
PD rect.	1.7*	0.8	1.1	3.0	1.7	0.7	3.0	4.0	4.6	2.1	5.9	5.7	5.5*	5.5*	7
T2 rect.	2.1*	0.8	1.6	3.5	1.6	0.8	4.3	5.3	4.2	2.9	5.5	5.3	5.3	5.3	7

The label "rect." indicates that the MR image was corrected for geometrical distortion before registration. See text for technique abbreviations. *One patient omitted. †Non-rigid transformations. ‡Results resubmitted after gold standard released. All errors are in units of mm.

TABLE 2
Maximum Errors for CT-to-MR Registration

MR Modality	Technique														N
	BA	CO	EL	HA	HE	HI‡	MAI	MAL	NO†	PE	RO1	RO2	RO3	RO4	
T1	6.4	6.7	6.0	51.8	11.0	2.8	12.8	61.4	10.4	7.3	26.0	21.8	17.8	18.8	7
PD	6.9	3.6	6.6	49.6	10.4	4.1	19.0	59.0	13.9	4.3	25.9	22.2	24.0	20.1	7
T2	9.1	3.4	4.1	50.6	13.6	4.2	6.3	59.5	9.7	7.2	26.7	22.0	19.4	19.8	7
T1 rect.	5.8*	3.8	2.6	48.2	2.1	2.3	14.2	60.9	9.6	5.9	27.8	22.2	18.1	18.2	6
PD rect.	5.9*	2.5	5.3	45.9	3.7	2.3	9.9	62.7	11.5	4.6	27.5	22.1	24.9*	20.2*	7
T2 rect.	7.4*	4.3	5.2	49.1	14.3	3.0	6.5	63.2	10.2	9.0	27.1	22.5	19.9	21.6	7

See notes in Table 1.

TABLE 3

MEDIAN ERRORS FOR PET-TO-MR REGISTRATION

MR Modality	Technique												N
	BA	CO	HA	HI‡	MAI	MAL	NO†	PE	RO3	RO4	WO1	WO2	
T1	4.6	3.6	2.8	3.2	3.5	4.2	3.6	2.9	4.0	3.4	2.3	3.1	7
PD	5.2	2.9	4.2	3.1	4.7	4.0	4.1	3.3	4.3	3.3	2.9	3.1	7
T2	4.7	2.8	2.7	2.4	5.3	4.9	4.6	3.3	4.0	3.6	3.6	3.4	7
T1 rect.	3.2	2.8*	3.6	2.5	3.9	3.6	3.9	2.8	3.8	3.6	2.0	2.0	4
PD rect.	4.5	3.0	3.2	3.0	4.7	3.6	4.4	2.8	3.6*	4.1*	2.5	2.3	5
T2 rect.	3.9	2.0	3.3	2.2	4.0	3.6	5.2	2.9	3.8	3.4	2.5	2.4	5

See notes in Table 1.

TABLE 4

Maximum Errors for PET-to-MR Registration

MR Modality	Technique												N
	BA	CO	HA	HI[‡]	MAI	MAL	NO[†]	PE	RO3	RO4	WO1	WO2	
T1	11.5	12.7	12.1	9.3	10.6	8.5	11.4	10.0	9.4	5.9	5.8	6.0	7
PD	11.2	9.2	10.3	8.1	9.8	9.3	8.9	11.3	8.8	7.1	6.9	6.3	7
T2	12.3	7.5	17.4	8.3	15.0	12.3	7.3	13.4	9.0	7.3	8.4	7.3	7
T1 rect.	6.0	3.7*	17.7	6.0	7.7	8.4	14.2	7.9	7.3	8.9	4.2	5.0	4
PD rect.	11.0	7.3	10.1	7.5	9.2	9.4	7.4	11.0	6.6*	6.6*	5.5	5.4	5
T2 rect.	9.8	7.1	10.2	9.3	10.9	12.4	11.2	15.2	5.8	7.1	6.0	6.1	5

See notes in Table 1.

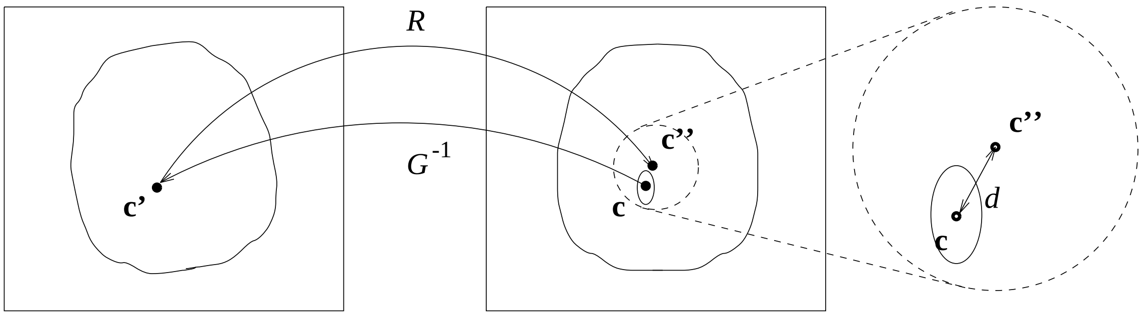

Fig. 1. Calculation of the accuracy of a retrospective registration at a VOI. A VOI (represented here by an ellipse) is defined in the "To" modality (right image). The centroid voxel of the VOI is found and converted from a voxel index to a millimetric position $\mathbf{c}$ using the known voxel size for the image volume. The inverse of the gold-standard rigid-body transformation G is applied to the point $\mathbf{c}$, giving point $\mathbf{c}' = G^{-1}(\mathbf{c})$ in the "From" modality (left image). Then, the retrospective transformation R is applied to $\mathbf{c}'$, giving point $\mathbf{c}''$. The registration error of the retrospective transformation at the centroid of the VOI is taken to be the Euclidean distance d between the points $\mathbf{c}$ and $\mathbf{c}'' = R(G^{-1}(\mathbf{c}))$.

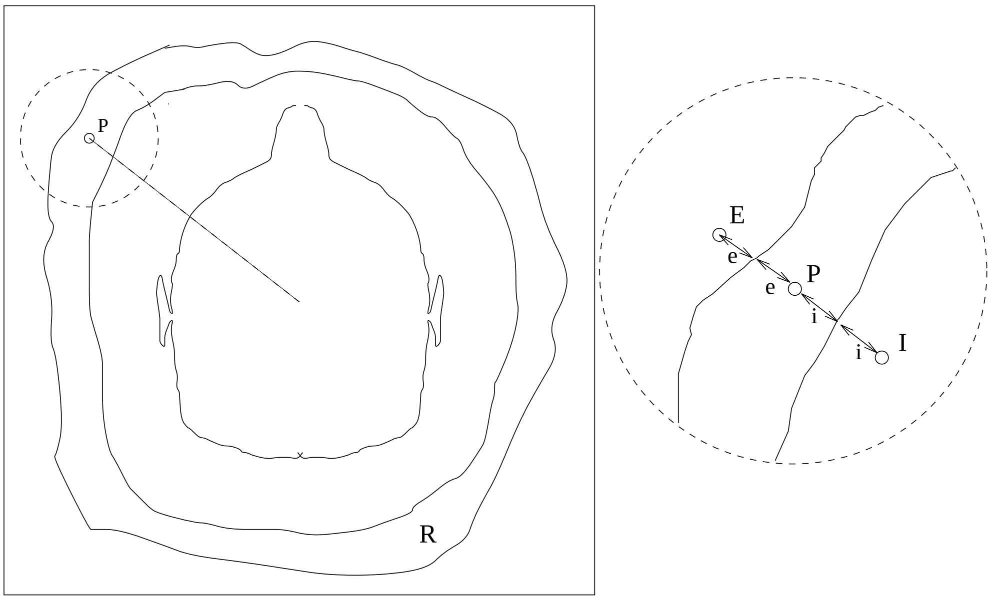

Fig. 2. Calculation of points for interpolation of background patterns in CT and PET images. See text for an explanation.

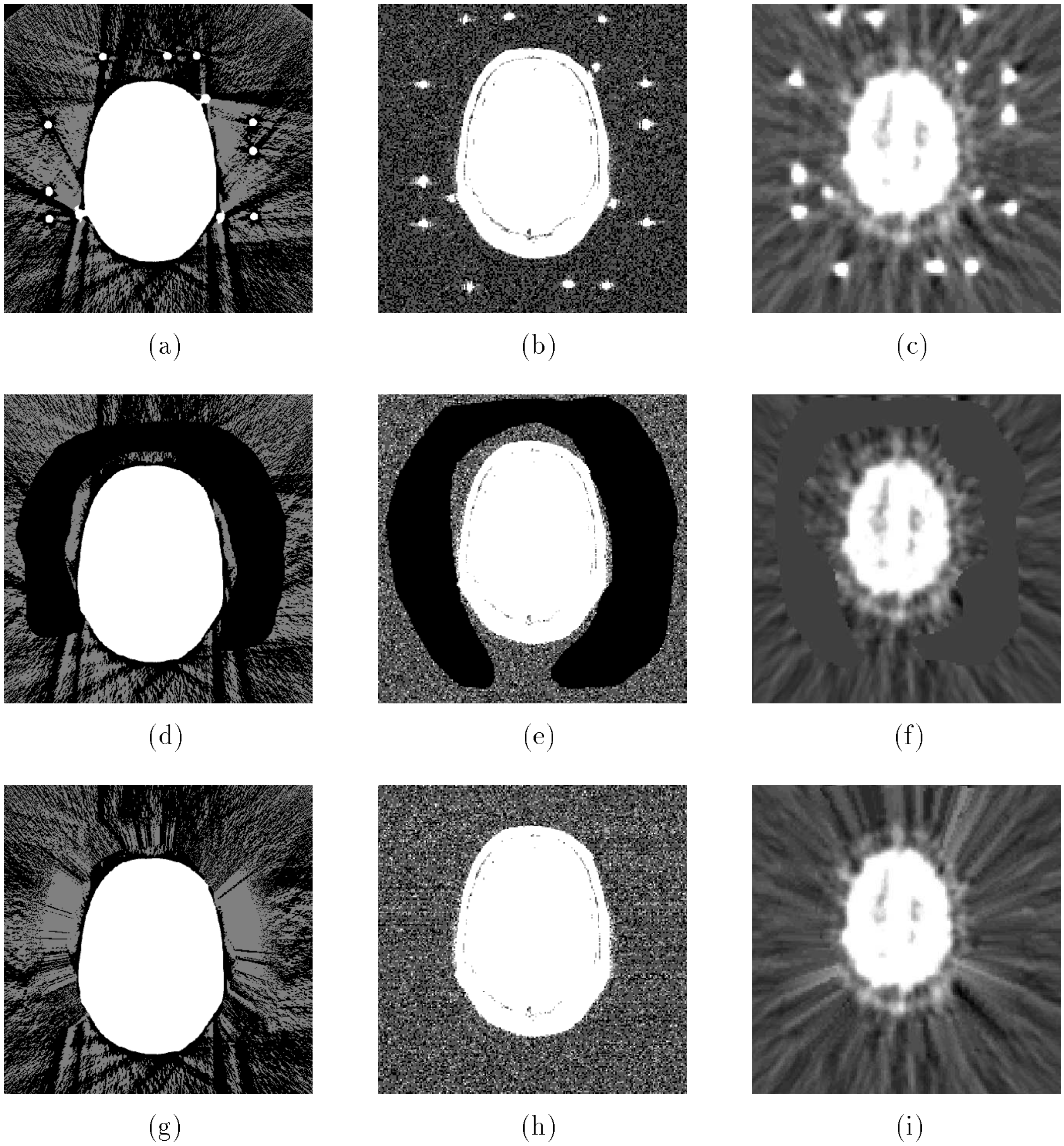

Fig. 3. Removal of fiducial markers and stereotactic frame. The top row (*a–c*) shows sample original image slices from CT (*a*), MR (*b*), and PET (*c*). The stereotactic frame (bright circular spots towards the edge of the image) and fiducial markers (three bright spots near the head) are clearly visible in all three modalities. The window and level have been set to show the background artifacts. The middle row (*d–f*) shows the same image slices after the region R has been outlined and zeroed. All trace of the stereotactic frame and fiducial markers has been removed. The bottom row (*g–i*) shows the image slices after reconstruction of the background in the region R. For MR, the replaced area is indistinguishable from the rest of the background. For CT and PET, there are slight discontinuities in the direction of the stripes, but the intensity changes relatively smoothly.

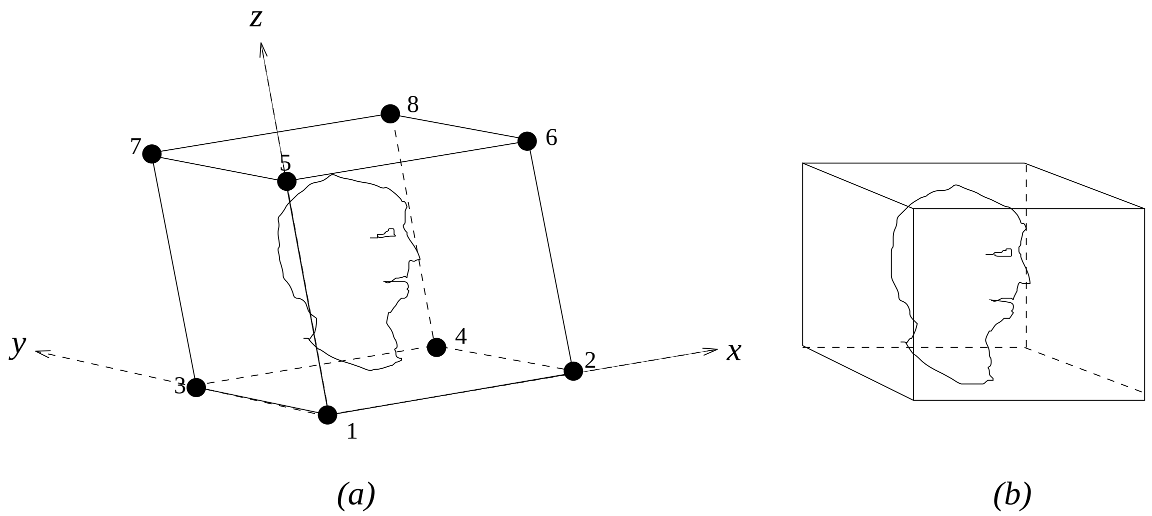

Fig. 4. Image volumes prior to registration: (*a*) "From" volume and (*b*) "To" volume. The points 1 to 8 are defined as the corners of the volume, relative to the axes shown, and form the "From" columns of the registration table.

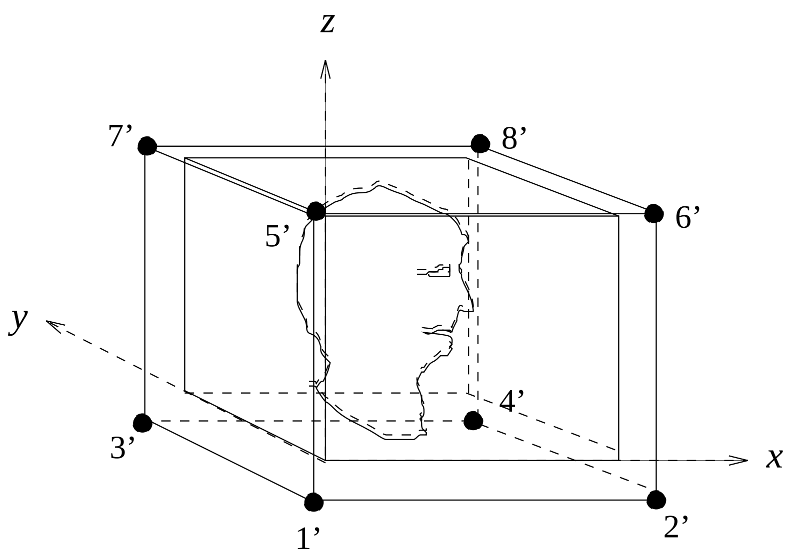

Fig. 5. Image volumes after registration, so that the heads in the two volumes are almost perfectly aligned with each other. The points $1'$ to $8'$ are used to specify the transformation and form the “To” columns of the transformation table.