Retrospective Intermodality Registration Techniques for Images of the Head: Surface-Based Versus Volume-Based

Jay West, J. Michael Fitzpatrick,* *Member, IEEE*, Matthew Y. Wang, Benoit M. Dawant, *Member, IEEE*, Calvin R. Maurer, Jr., *Member, IEEE*, Robert M. Kessler, and Robert J. Maciunas

***Abstract*—The primary objective of this study is to perform a blinded evaluation of two groups of retrospective image registration techniques, using as a gold standard a prospective marker-based registration method, and to compare the performance of one group with the other. These techniques have already been evaluated individually [27]. In this paper, however, we find that by grouping the techniques as volume based or surface based, we can make some interesting conclusions which were not visible in the earlier study. In order to ensure blindness, all retrospective registrations were performed by participants who had no knowledge of the gold-standard results until after their results had been submitted. Image volumes of three modalities: X-ray computed tomography (CT), magnetic resonance (MR), and positron emission tomography (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 (VOI's). Our results indicate that the volume-based techniques in this study tended to give substantially more accurate and reliable results than the surface-based ones for the CT-to-MR registration tasks, and slightly more accurate results for the PET-to-MR tasks. Analysis of these results revealed that the rotational component of error was more pronounced for the surface-based group. It was also apparent that all of the registration techniques we examined have the potential to produce satisfactory results much of the time, but that visual inspection is necessary to guard against large errors.**

***Index Terms*—Computed tomography, image registration, magnetic resonance, positron emission tomography, surface-based versus volume-based.**

Manuscript received January 12, 1998; revised January 4, 1999. This work was supported in part by the National Institutes of Health under Project 1 R01 NS33926-01. Preliminary results of this work were published in the *Journal of Computer Assisted Tomography*. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was C. Roux. *Asterisk indicates corresponding author.*

J. West, *J. M. Fitzpatrick, and M. Y. Wang are with the Department of Computer Science, Vanderbilt University, 1500 21st Avenue South, Nashville, TN 37212 USA.

B. M. Dawant is with the Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN 37212 USA.

C. R. Maurer, Jr. is with the Department of Biomedical Engineering, Vanderbilt University, Nashville, TN 37212 USA.

R. M. Kessler is with the Department of Radiology, Vanderbilt University, Nashville, TN 37212 USA.

R. J. Maciunas is with the Department of Neurological Surgery, Vanderbilt University, Nashville, TN 37212 USA.

Publisher Item Identifier S 0278-0062(99)03154-7.

I. INTRODUCTION

We present here a new analysis of the previously published [27] results of a study.[1] Most of this analysis was first presented at the conference CVRMed/MRCAS 1997 in Grenoble, France [26]. By dividing the registration methods evaluated in this study into two groups, volume based and surface based, it is possible to gain more statistical power from the results. We label as volume based, any technique which performs registration by making use of a relationship between voxel intensities within the images and as surface-based, any technique which works by minimizing a distance measure between two corresponding surfaces in the images to be matched. We then compare the mean accuracy achieved on a set of registration tasks by the methods in the volume and surface groups.

There were two registration tasks evaluated in this study, X-ray computed tomography to magnetic resonance (CT to MR) and positron emission tomography to MR (PET to MR), and these tasks were broken into subtasks according to the type of MR and whether the MR image was corrected for geometrical distortion. The image data sets of nine patients were used, of which seven contained both CT and MR scans and seven others contained PET and MR scans. MR scans of three types: T1, proton density (PD), and T2 were included. 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 [3], [20]. The first step toward evaluation is the calculation of the answers, i.e., 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. This was done by manually segmenting the frame and markers out of the images and replacing the removed regions by a simulated background pattern in each imaging modality. We call this process air brushing.

These air-brushed image volumes were then made available on the Internet via file transfer protocol (FTP) so that each site could apply its own retrospective technique. The resulting transformations were reported by specification of the motion of the eight corner points of the volume. A measurement of error was made by calculating the error relative to the gold

[1] These results and those subsequent to the study are now available on the World Wide Web at http://cswww.vuse.vanderbilt.edu/~jayw/results.html.

standard over a set of specified regions known as volumes of interest (VOI's).

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.

II. METHODS

A. Image Acquisition

The CT images were acquired using a Siemens DR-H scanner, the MR images were acquired using a Siemens SP 1.5-Tesla scanner, and the PET images with a Siemens/CTI ECAT 933/08–16 scanner. The CT volumes have a resolution of 512 × 512 pixels in the x and y directions and contain 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. For MR, T1-weighted (T1), PD-weighted (PD), and T2-weighted (T2) spin-echo images were acquired. These volumes have a resolution of 256 × 256 pixels in the x and y directions and contain between 20 and 26 slices in the z direction. The voxel size is between 1.25 and 1.28 mm in the x and y directions and 4.0 mm in z. 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 a TE of 20 ms and TR of 2550 ms (20 slices) or 3000 ms (26 slices); and the T2 with a 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 min after injection and was continued for 25 min. The volumes have a resolution of 128 × 128 pixels in x and y and contain 15 slices in the z direction. The voxel size is 2.59 mm in x and y and 8.0 mm in z. 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 [3], [20]. 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

The markers are 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. Plastic marker bases or posts are screwed into the outer table of the skull of the patient. The markers are attached to the posts during 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 previous publications [20], [25].

We define the position of a marker as its centroid and call the determination of this position fiducial localization. We calculate an intensity-based centroid for each marker using the localization technique described in Wang *et al.* [25]. The large point-spread function increases the effective size of the marker substantially in PET. Because the marker is spread over more than one slice, it becomes possible to use intensity weighting in determining the centroid in the direction perpendicular to a slice to an accuracy better than the slice thickness. The published localization algorithm, therefore, was modified for this project to take advantage of the increased effective size for PET.

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 [19], [21]. We have implemented the closed-form solution developed by Arun *et al*. [1]. We define the fiducial registration error (FRE) as the root-mean-square (rms) distance between corresponding fiducials after registration and transformation.

This registration process 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. A slightly different procedure was used for each modality in order to reconstruct as accurately as possible the nature of the background artifacts in each type of image [27]. For each air-brushed image volume, an ASCII header was prepared. Each header contains a description of the image to which it relates, giving the pixel size, slice thickness, resolution, data format, and acquisition information. In the MR case, the headers also contain a line stating the image to be either an unaltered volume or a volume to which the distortion-corrections technique has been applied. The headers conform to the Interfile standard [2], [5].

F. Communications

After creation of the images and headers was complete, a login name, password, and Internet address were provided by e-mail to all project participants 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. These positions were provided via FTP by Vanderbilt for every CT and PET volume.

After the retrospective registration transformation was determined, the transformed positions of these eight points relative to the origin of the To modality were computed by each site. The field of view of the two volumes is typically different, so it is important to specify which volume provides the origin relative to which the transformed positions are calculated. A more detailed account of the communication of transforms is given in the first publication related to this study [27].

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 set of transformed positions 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. This protocol is thus restricted to evaluation of rigid-body transformations.

Each transformation was transmitted to Vanderbilt by e-mailing an ASCII file containing both the original and transformed points.

G. 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 to PET-to-MR registration and some were suitable for both cases.

We have, for the purposes of this paper, divided the registration techniques into two categories. Surface-based registrations were performed by Barillot and Lemoine [11]; Harkness [22]; Hemler *et al*. [7] (CT-to-MR only); Pelizzari [22]; and Robb and Hanson [9], [10] (four techniques for CT-to-MR, two of which were also applied in PET-to-MR). Volume-based registrations were performed by Collignon *et al*. [4]; Van den Elsen *et al*. [24] (CT-to-MR only); Hill and Studholme [23]; Maintz *et al* [13], [14], [15]; and Woods [28] (two techniques, PET-to-MR only). Two sites which contributed registrations to the original study are not included here. The technique of Noz *et al.* [12] was omitted as it performed nonlinear transformations and that of Malandain and Pennec [16]–[18] was also omitted as it is best described as a hybrid between volume- and surface-based registration and, thus, could not be unambiguously placed in either of the groups used. All the volume-based techniques, and all but one of the surface-based techniques (that of Barillot and Lemoine), used in this study were fully automated in terms of the registration step, i.e., they did not require that the images be manually placed in close alignment before registration took place.

H. Data Analysis

At Vanderbilt, after the transformations were received from each site and the corresponding rigid-body transformations had been determined, the next step in the evaluation is to perform a comparison between these registrations and the fiducial-based ones. In some cases, errors were discovered by the remote sites after the transformations had been submitted and the gold-standard registrations had been distributed to the investigators at these sites. We allowed a modification of the results submitted by Hill *et al.*, and we removed erroneous transformations from the sets being evaluated for Barillot *et al.*, Collignon *et al.*, and Robb *et al*. A full discussion of these cases is given in [27].

In collaboration with a neurological and a neurosurgical expert, a set of VOI's was chosen that represents areas of neurological and/or surgical interest. Each VOI was manually segmented within one of the MR image volumes. This procedure was repeated for each patient data set used. The VOI's 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. The centroid pixel of the VOI is found and its position is converted from a voxel index to a millimetric position $\boldsymbol{c}$ in the To modality, using the known voxel size for the image volume (Fig. 1).

Let $\boldsymbol{R}_g$ and $\boldsymbol{t}_g$ be the rotation matrix and translation vector, respectively, of the gold-standard rigid-body transformation G and let $\boldsymbol{R}_r$ and $\boldsymbol{t}_r$ be the rotation and translation components of the retrospective transformation R. The point $\boldsymbol{c}'$ in the From modality is defined so that $\boldsymbol{c}$ is the mapping of $\boldsymbol{c}'$ under the gold-standard transformation. Thus

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

By inverting (1) we obtain

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

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

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

The discrepancy between the registered target position of the retrospective method and that of the gold standard is $\boldsymbol{d} = \boldsymbol{c}'' - \boldsymbol{c}$. In image registration, the target registration error (TRE) at a given point is the distance between that point and the corresponding location in the other image after registration

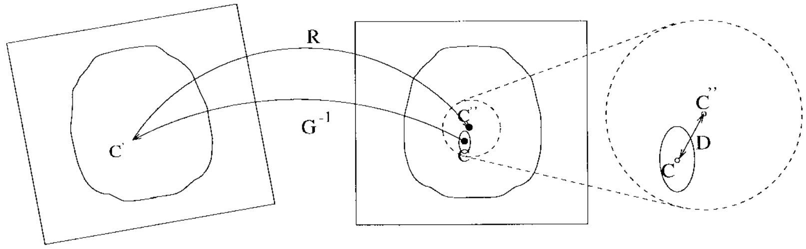

Fig. 1. Calculation of the accuracy of a retrospective registration at a VOI. A VOI (represented here by an ellipse) is defined in the To modality (right image). The centroid voxel of the VOI is found and converted from a voxel index to a millimetric position $\boldsymbol{c}$ using the known voxel size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

has been performed. We define the TRE of the retrospective registration at the anatomic location of the VOI to be the Euclidean distance d between $\boldsymbol{c}$ and $\boldsymbol{c}''$, i.e., TRE $= d = ||\boldsymbol{d}||$.

The particular anatomic positions corresponding to the VOI's 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, there appeared to be no consistent pattern governing the variation of TRE among the VOI's. For this reason, we are reporting statistics only for the pooled VOI's.

The results are divided into two groups:

- those from volume-based registration methods, i.e., those methods which make use of a relationship between voxel intensity values in three-dimensional (3-D) regions of the two image volumes to be matched;
- those from surface-based methods, i.e., those which perform registration by minimizing a distance measure between two corresponding surfaces, one of which has been segmented or derived in each image volume.

For each modality pair e.g., (CT-T1), a volume-based and surface-based registration error for each patient was derived. This was done by taking the mean of the list of errors over all VOI's and all techniques categorized as volume or surface, respectively. As a measure of the performance of the two groups, in each case the mean error over all patients was taken for each modality pair. This is presented in Table I. We also present the percentage of instances in which a member of each group had a registration error of more than 10 mm. One such instance is defined as occurring, for the task of registering a particular modality pair, when for one patient, one of the techniques of the group has an error greater than 10 mm, as measured at one VOI.

First, we investigated whether there was any significant difference between the results of the volume- and surface-based registrations. This was done for each modality pair by a paired comparison of the mean registration error for each patient. For CT-to-MR registration, in half the cases (CT-to-T2, CT-to-PD rectified, and CT-to-T2 rectified) the volume-based techniques are significantly more accurate (two-tailed paired t test $P < 0.05$). In the remaining three cases (CT-to-PD, CT-to-T1, and CT-to-T1 rectified) volume-based techniques are more accurate with a marginal significance ($P < 0.10$). For PET-to-MR registration, the volume-based methods are significantly more accurate in the PET-to-PD case ($P < 0.05$). Otherwise, no significant difference in accuracy was observed between the two groups.

Second, it was our goal to find out whether MR distortion correction had any effect on the accuracy of the retrospective registration methods. We performed another paired comparison, this time between the results of a registration type on each patient for registering CT and PET to an MR modality, with and without distortion correction having been applied. For the surface-based group no significant changes in accuracy were seen. In the volume-based group, however, CT-to-PD rectified registration is significantly more accurate (two-tailed paired t test $P < 0.05$) than CT-to-PD. CT-to-T1 rectified registration is more accurate than CT-to-T1 with a marginal significance ($P < 0.10$). Distortion correction did not have any significant effect on the PET-to-MR results.

For both these sets of conclusions, however, it must be noted that we could not take account of the intravariation within the samples used to derive the mean values for the comparison, as the number of patient data sets was too small. In other words we assume the measurements are independent even though they came from the same patient. It is thus possible that the significances we report here are overestimates.

Finally, we looked at the translation and rotation components of the registration errors for each group. It was our goal here to determine whether there was a different tendency toward rotational or translational error for the two groups. For each retrospective transformation R we computed the error transformation $E = RG^{-1}$ where G is the gold-standard transformation for that registration task. We took the ratio of the rotation angle (in radians) of E to the magnitude of the translation component of E (in millimeters), measured at the centroid of the volume. We compared separately the results for CT-to-MR and PET-to-MR registration. The number of transformations forming each group was 311 and 185, respectively, for the surface- and volume-based groups and CT-to-MR registration and 167 and 174 for the surface- and volume-based groups and PET-to-MR registration.

TABLE I
REGISTRATION ERRORS

Modality Pair	Surface group Mean (std.)	Surface group % > 10	Volume group Mean (std.)	Volume group % > 10	N
CT-T1	5.7 (7.8)	11.7	2.9 (2.4)	1.2	7
CT-PD	5.8 (8.0)	11.3	2.9 (2.5)	1.6	7
CT-T2	6.3 (7.9)	12.3	2.4 (1.4)	0.0	7
CT-T1 rect.	6.1 (8.3)	13.1	2.0 (2.5)	1.9	6
CT-PD rect.	5.7 (7.8)	12.0	1.8 (2.0)	0.0	7
CT-T2 rect.	6.1 (7.6)	12.1	2.1 (1.6)	0.0	7
PET-T1	3.9 (2.0)	1.3	3.5 (2.1)	2.3	7
PET-PD	4.4 (2.1)	2.3	3.6 (1.9)	0.0	7
PET-T2	4.3 (2.6)	2.9	4.0 (2.7)	4.5	7
PET-T1 rect.	3.9 (2.3)	1.6	2.7 (1.4)	0.0	4
PET-PD rect.	3.9 (2.0)	3.2	3.5 (1.7)	0.0	5
PET-T2 rect.	3.9 (2.3)	2.1	3.5 (2.4)	3.8	5

TABLE II
MEAN ROTATIONAL (R), TRANSLATIONAL (T), AND RATIO ROTATIONAL/TRANSLATIONAL (R/T) ERROR FOR CT-MR REGISTRATION

Group	R (± sd) (radians)	T (± sd) (mm)	R/T (± sd) (radians/mm)
Surface-based	0.0409 (0.0337)	2.1069 (1.1088)	0.0213 (0.0152)
Volume-based	0.0222 (0.0269)	1.3859 (0.5405)	0.0151 (0.0118)

TABLE III
MEAN ROTATIONAL (R), TRANSLATIONAL (T), AND RATIO ROTATIONAL/TRANSLATIONAL (R/T) ERROR FOR PET-MR REGISTRATION

Group	R (± sd) (radians)	T (± sd) (mm)	R/T (± sd) (radians/mm)
Surface-based	0.0473 (0.0322)	1.8444 (0.4208)	0.0255 (0.0149)
Volume-based	0.0372 (0.0228)	1.6765 (0.4902)	0.0229 (0.0140)

In Tables II and III we show the mean translation, rotation, and ratio translation/rotation errors for each group. The numbers in parentheses give the standard deviation for each group. Translation was observed at the centroid of the volume. We found, using a two-tailed t test assuming unequal variance, that the translation and rotation components of error were each significantly greater for the surface-based group than for the volume-based group ($P < 0.05$) in both CT-to-MR and PET-to-MR registration. We also found that the ratio of rotational to translational error was significantly greater for the surface-based group in CT-to-MR registration ($P < 0.05$). For PET-to-MR registration, this ratio was greater for the surface-based group to a marginal significance ($P < 0.10$).

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 MRI (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.

The validity of these evaluations depends on the accuracy of the gold-standard registrations. If the transformation errors of the fiducial marker and a retrospective technique are uncorrelated, the following simple relationship holds for the rms of the observed TRE, the true TRE (TRE_t), and the gold-standard TRE (TRE_g):

$$\mathrm{rms}[\mathrm{TRE}] = \sqrt{\mathrm{rms}^2[\mathrm{TRE}_t] + \mathrm{rms}^2[\mathrm{TRE}_g]}. \tag{4}$$

While we cannot measure $\mathrm{rms}[\mathrm{TRE}_g]$ directly, we can estimate it using numerical simulations [19], [21]. According to the simulation results, 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. Using the estimation method described in [19] and [21] we then have that $\mathrm{rms}[\mathrm{TRE}_g]$ is approximately 0.39 mm for CT-to-MR and 1.65 mm for PET-to-MR. The larger TRE's for registrations involving PET are to be expected because of the larger voxels in this modality.

These simulations assume there is no geometrical distortion. Thus, these simulation results apply only to the registrations obtained using MR images corrected for geometrical distortion, assuming that any remaining distortions are uncorrelated. The gold-standard TRE's can be expected to be somewhat larger for the uncorrected MR images.

If a retrospective technique's accuracy is approximately the same as the gold standard, then it follows from (4) that the rms of the observed TRE will be approximately $\sqrt{2}\ \mathrm{rms}[\mathrm{TRE}_g]$, which is approximately 0.55 and 2.33 mm for the CT-to-MR and PET-to-MR cases, respectively. The smallest rms values in this study (taken for any modality pair and any technique, but averaged over all patients) are 1.0 mm for CT-to-MR and 2.3 mm for PET-to-MR. Both these values were achieved by techniques in the group classified as volume-based registration. These small values suggest that the accuracy of some of the volume-based retrospective techniques approaches the accuracy of the bone-implanted fiducial marker method, at least for PET-to-MR registration.

The column headed $\% > 10$ in Table I gives the percentage of instances in which the measured error is greater than 10 mm. This column is included to provide information about very large misregistration errors. As can be seen from this column, for the case of surface-based registration of CT to MR, there were errors of at least 10 mm in at least 11% of the instances recorded, regardless of the type of MR acquisition, indicating that the quality of surface-based CT-to-MR registration should be checked visually. Such large errors are less common both for surface-based registration of PET to MR and for all volume-based registrations. There are six cases (entries of 0.0) in which the maximum observed errors are less than 10 mm. The maximum observed errors for these six cases (not tabulated) range from 6.3 to 9.9 mm. As with the smaller means and standard deviations, these smaller percentages are an indication of greater registration success, but the presence of any large errors suggests that visual inspection is an important adjunct for these techniques as well.

In Section III we showed that the errors in the surface-based group had a significantly larger rotational component than those in the volume-based group. We hypothesize that the reason for this is that, in cases where the segmented surface is close to symmetric about some axis, the distance measures minimized by surface-based methods may often be insensitive to rotations about this axis. Because the contents of the head generally have less rotational symmetry than the surfaces of the scalp and brain, volume-based methods may be less prone to rotational errors than surface-based methods for images of the head.

V. Conclusion

Ten groups of investigators applied 14 techniques, of which six were volume-based and eight were surface-based, to selected registration tasks involving the registration of CT and/or PET to MR. Our results indicate that, on the patient data sets we provided, the volume-based registration techniques in the study tended to be significantly or marginally significantly more accurate than the surface-based techniques in the CT-to-MR case. In the PET-to-MR case the volume-based methods were slightly more accurate, but the difference was not significant except for the PET-to-PD modality pair. By grouping the techniques in this way, we have shown that we have sufficient statistical power to make these interesting conclusions. This was not possible when the techniques were evaluated individually [27]. Our analysis of these results revealed that for CT-to-MR registration the rotational component of error was more pronounced for the surface-based group: a similar but less marked effect was observed for PET-to-MR registration. It is also apparent that all of the retrospective methods in this study have the potential to produce satisfactory results much of the time, but that visual inspection is necessary to guard against large errors. It would usually be possible to perform such an inspection for a single technique by a simple extension of the software performing the registration. However, all that is necessary for visual inspection is a knowledge of the transformation output by the registration algorithm. Visual inspection may be performed using simply this transformation, the images themselves, and software capable of reformatting one image according to the registration transformation so that it may be directly compared with the other [8], [6].

Acknowledgment

The authors would like to thank Dr. Y. Shyr for the time and effort he spent advising them about the statistical methods used in this study, Dr. J. Votaw for answering many questions about PET imaging, and S. Gadamsetty for helping to collect and process patient images.

References

[1] 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, Sept. 1987.

[2] B. S. Baxter, L. E. Hitchner, and G. Q. Maguire, Jr., "A standard format for digital image exchange," *Am. Assoc. Physicists Med.*, 1982.

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

[4] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Maral, "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. Dordrecht, The Netherlands: Kluwer, 1995, pp. 263–274.

[5] 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.

[6] J. M. Fitzpatrick, D. L. G. Hill, Y. Shyr, J. West, C. Studholme, and C. R. Maurer, Jr., "Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain," *IEEE Trans. Med. Imag.*, vol. 17, pp. 571–585, Aug. 1998.

[7] 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.

[8] J. J. Jacq and C. Roux, "A direct multi-volume rendering method application to visual assessment of 3-D image registration algorithms," in *Visualization in Biomedical Computing. Lecture Notes in Computer Science 1131*, K. H. Hoehne and R. Kikinis, Eds. Berlin, Germany: Springer–Verlag, 1996, pp. 53–62.

[9] 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. 1660, pp. 356–366, 1992.

[10] 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. 1808, pp. 196–213, 1992.

[11] 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. 2164, pp. 46–56, 1994.

[12] 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, Mar. 1991.

[13] 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. Berlin, Germany: Springer-Verlag, 1995, pp. 219–228.

[14] ——, "Comparison of edge-based and ridge-based registration of CT and MR brain images," *Med. Image Anal.*, vol. 1, pp. 151–161, 1996.

[15] ——, "Evaluation of ridge seeking operators for multimodality medical image matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, pp. 353–365, Apr. 1996.

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

[17] ——, "Rigid registration of 3-D objects by motion analysis," in *Proc. 12th Int. Conf. Pattern Recognition*, 1994, pp. 579–581.

[18] ——, "Physically based rigid registration of 3-D free-form objects: Application to medical imaging," INRIA, Sophia Antipolis Cedex, France, Tech. Rep. 2453, Jan. 1995.

[19] 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. 1652, pp. 271–282, 1992.

[20] 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.

[21] 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. Berlin, Germany: Springer-Verlag, 1995, pp. 1197–1202.

[22] 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.

[23] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated registration of truncated MR and CT datasets of the head," *Proc. Br. Mach. Vision Assoc.*, 1995, pp. 27–36.

[24] 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. 2359, pp. 227–237, 1994.

[25] 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, June 1996.

[26] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, and R. J. Maciunas, "Retrospective intermodality registration techniques: Surface-based versus volume-based," in *CVRMed-MRCAS '97*, J. Troccaz, E. Grimson, and R. Mösges, Eds. Berlin: Springer-Verlag, 1997, pp. 151–160.

[27] 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, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, 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," *J. Comput. Assist. Tomogr.*, vol. 21, pp. 554–566, 1997.

[28] 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.