Journal of Computer Assisted Tomography
18(6):954–962, November/December

Quantitative Comparisons of Image Registration Techniques Based on High-Resolution MRI of the Brain

Stephen C. Strother, Jon R. Anderson, Xiao-Liang Xu, Jeih-San Liow, David C. Bonar, and David A. Rottenberg

Objective: A variety of methods for matching intrasubject MRI-MRI, PET-PET, or MRI-PET image pairs have been proposed. Based on the rigid body transformations needed to align pairs of high-resolution MRI scans and/or simulated PET scans (derived from these MRI scans), we obtained general comparisons of four intrasubject image registration techniques: Talairach coordinates, head and hat, equivalent internal points, and ratio image uniformity. In addition, we obtained a comparison of stereotaxic Z frames with a customized head mold for MRI-MRI image pairs.

Materials and Methods and Results: Each technique was quantitatively evaluated using the mean and maximum voxel registration errors for matched voxel pairs within the brain volumes being registered.

Conclusion: We conclude that fiducial markers such as stereotaxic Z frames that are not rigidly fixed to a patient's skull are inaccurate compared with other registration techniques, Talairach coordinate transformations provide surprisingly good registration, and minimizing the variance of MRI-MRI, PET-PET, or MRI-PET ratio images provides significantly better registration than all other techniques tested. Registration optimization based on measurement of the similarity of spatial distributions of voxel values is superior to techniques that do not use such information.

Index Terms: Image registration—Image quality—Emission computed tomography—Magnetic resonance imaging.

Medical image registration or "image fusion" represents a specific application within the field of close-range photogrammetry (1). The goal of medical image registration is to geometrically align two or more image volumes so that voxels representing the same underlying anatomical structure may be superimposed. A pair of image volumes might be from the same subject and the same modality (i.e., intrasubject and intramodality registration), from the same subject but different modalities (i.e., intrasubject and intermodality), or from the same modality but different subjects (i.e., intersubject) (2). Registration techniques are typically used to ensure that the same anatomical regions of the brain are being sampled in voxel-based data analysis, e.g., [^{15}O] water activation–subtraction studies (3,4). Many registration techniques have been proposed in recent years, and they have typically been evaluated in isolation or against one other technique. Moreover, it is difficult to understand the relative significance of reported registration errors involving axis-specific translations and rotations: Is a 1° error about the x-axis better or worse than a 1 mm translation error along the y-axis?

We previously attempted to overcome some of these difficulties in comparing published registration techniques by using a consistent summary measure of mean and maximum misregistration within real brain image volumes (5). Our goal was to obtain lower error bounds for image registration techniques by comparing them under conditions expected to provide near optimal registration performance. For a single observer, we found (5) that for comparisons of original and mathematically trans-

From the PET Imaging Service, VA Medical Center (S. C. Strother, J. R. Anderson, X.-L. Xu, J.-S. Liow, D. C. Bonar, and D. A. Rottenberg), and Departments of Neurology (S. C. Strother, J. R. Anderson, X.-L. Xu, and D. A. Rottenberg) and Radiology (S. C. Strother, J.-S. Liow, and D. A. Rottenberg), University of Minnesota, Minneapolis, Minnesota, U.S.A. Address correspondence and reprint requests to Dr. S. C. Strother at PET Imaging Service (11P), 1 Veterans Dr., VA Medical Center, Minneapolis, MN 55417, U.S.A.

formed MRI image pairs, the principal axis (PA) technique (6) provided similar mean registration errors (≈1 mm) to those from head and hat (HH) (7) and equivalent internal points (EIP) (8), with larger mean errors (≈2 mm) from Talairach coordinates (TA) (3,9); with MRI and simulated PET image pairs, EIP (mean errors ≈1.5 mm) slightly outperformed HH and PA (mean errors ≈2.2 mm), with larger mean errors from TA (≈3.1 mm). Comparison of original with repeated and, in addition, deliberately mispositioned MRI scans produced estimated errors in the 1 to 2 mm range.

We have extended this preliminary intrasubject alignment data by performing a new study using two observers to compare interobserver and intertechnique registration errors for the ratio image uniformity technique (RIU) (2,10–12) compared with HH, EIP, and an improved TA procedure. In addition, more realistic simulated PET volumes, which incorporate reconstruction and noise effects, were created and used to examine PET-PET alignment for RIU, HH, EIP, and TA. We also tested ideal results for a stereotaxic Z frame (SZF) with a customized head mold (13,14) to separate algorithmic errors from those due to repositioning subjects' heads within the head molds.

MATERIALS AND METHODS

Transformation Matrices Using Homogeneous Coordinates

Our measurements of registration accuracy were based on 4 × 4 homogeneous transformation matrices (TMs) (15) that contained the scaling, rotation, and translation parameters necessary to match the voxels in one volume with the equivalent voxels in another volume. As noted by Ende et al. (16), homogeneous coordinate transformations are attractive because they are easy to implement, the inverse transform is simply the inverse of the original transformation matrix, successive transformations (e.g., rotations about different axes) may be implemented as repeated matrix multiplications, and they allow nonrigid body skew and perspective transformations similar to the skew and asymmetry outlined by Greitz et al. (17,18). In this article we deal only with rigid body transformations involving translation, rotation, and limited scaling; the scaling factors are used to compensate for different voxel sizes. As an intermediate scaling step to obtain more accurate matches between brain surfaces from different modalities, the HH scaling factors were allowed to vary (19,20). However, the final TMs derived for the registration techniques (RIU, HH, EIP, TA, and SZF) used common scaling parameters based on the voxel sizes of the input image volumes. The techniques differ in their use of different algorithms to generate the translation and rotation parameters.

MRI Scans and Simulated PET Scans

Each MRI scan was obtained from a GE Signa scanner and encompassed the entire axial extent of the brain (~50 slices). Scans were composed of two SE volumes (1 × 1 × 3 mm voxels) acquired with TEs of 20 ms (early) and 80 ms (late). Eight normal volunteers (aged 27–38 years) participated in our study, but repositioned headholder scans were obtained only in six, with interscan intervals from 1 to 86 weeks. The two scans were performed with a headholder consisting of a customized foam mold (5) rigidly fixed in an acrylic frame bearing Z-shaped fiducial markers ("headholder MRI scan") and the customized headholder rotated 15–30° and repositioned axially 1–3 cm relative to the "normal" position of the previous two scans ("repositioned MRI scan").

Six mathematically misaligned MRI scans ("transformed MRI scans") were formed from the six headholder MRI scans. One TM was generated for each of the six subjects by randomly selecting rotations (<30°) and translations (<2.5 cm) for each axis. Six transformed MRI scans were then created by applying the TMs, followed by trilinear interpolation to the headholder MRI scans.

Six simulated PET scans were formed from the six headholder MRI scans using the resolution parameters and reconstruction algorithm of the CTI-Siemens 953B PET scanner (21). For each headholder MRI scan, we created gray matter (GM), white matter (WM), and CSF volumes from the two SE volumes using the segmentation technique described by Bonar et al. (22); formed a new volume by combining the GM, WM, and CSF volumes with a gray/white ratio that produced a good match between the slice means and frequency histograms of the simulated PET and a coplanar real PET scan; scaled the new volume so that the slice with the most counts had 650,000 total counts; forward-projected and smoothed (5.8 mm FWHM Gaussian) the scaled volume on a slice-by-slice basis; applied attenuation factors based on $\mu = 0.096\ cm^{-1}$ and then added Poisson noise; applied attenuation corrections and reconstructed 2D images (128 × 128) on a slice-by-slice basis forming a 128 × 128 × 50 slice volume; axially smoothed the volume with a 5 mm FWHM Gaussian kernel and downsampled axially to form a simulated PET volume of 128 × 128 × 31 slices (3.125 × 3.125 × 3.375 mm voxels); and applied the TM used in generating the corresponding transformed MRI volume, followed by trilinear interpolation, to the original simulated PET volume to create a "transformed PET volume."

Registration techniques were tested by two ob-

servers with experience in locating the anterior commissure (AC), posterior commissure (PC), and 24 "equivalent internal points" but without specialized neuroanatomical training. They registered headholder MRI and transformed MRI scans, headholder MRI and simulated PET scans, simulated PET and transformed PET scans, and headholder MRI and repositioned MRI scans.

Performance Measures

To measure the performance of registration techniques, we used mean (±SD) and maximum voxel registration errors (VREs) calculated for all brain voxel locations within the reference volume of each pair. The VRE was defined as the geometric distance in millimeters between the centers of anatomically equivalent voxels in a pair of image volumes after the application of a particular registration technique. For each registration technique, we calculated the group mean (±SD) of the individual mean and maximum VREs from the six registered volume pairs.

For the headholder MRI with transformed MRI or simulated PET and the simulated PET with transformed PET volumes, we calculated six residual TMs (one for each subject) by subtracting the estimated TM from the true TM; applied the residual TMs to the reference headholder MRI scans, creating six VRE volumes; and for each VRE volume calculated the mean (±SD) and maximum VREs for all brain voxels.

Since the true TMs for headholder and repositioned MRI-MRI scan pairs were unknown, mean (±SD) and maximum VREs were estimated for each pair based on correlation coefficients (5) derived using only intracranial voxels [the coefficients reported in ref. 5 were calculated using all image voxels]. We calculated the correlation coefficient and mean (±SD) and maximum VREs for headholder and transformed MRI pairs and used a quadratic regression to predict mean (±SD) and maximum VREs from the correlation coefficients. These quadratic functions were then used to estimate the VREs of headholder and repositioned MRI-MRI scan pairs based on their correlation coefficients after registration.

Stereotactic Z Frame (SZF)

For the SZF technique, subjects were scanned in customized foam head molds rigidly fixed in an acrylic frame containing Z-shaped fiducial markers filled with copper sulfate solution. The alignment procedure is briefly described by Anderson et al. (5).

Talairach Coordinates (TA)

The TA technique outlined by Anderson et al. (5) was a direct implementation of the approach described by Friston et al. (23), which was modified for this study. The locations of the ventral aspect of the anterior corpus callosum, the ventral aspect of the thalamic nuclei, the ventral aspect of the posterior corpus callosum, and the occipital pole were identified. Each location was moved inferiorly by an amount equal to its mean distance from the AC-PC line in eight headholder MRI scans; the estimated AC-PC line was determined by performing a weighted linear regression on the four adjusted locations. This modified approach is similar to the two stage fitting procedure suggested by Evans et al. (24). Once the midsagittal plane and the commissures were identified for each volume, translation parameters required to align the AC points were calculated and applied. Finally, the rotations needed to align the AC-PC lines and the midsagittal planes were computed so that the Talairach coordinate systems for the two image volumes coincided.

Head and Hat (HH)

Extrameningeal tissues and air were "stripped" from each MRI slice by applying a region-growing algorithm (22) and then detecting the boundaries of the resulting "brain masks" of 0's (nonbrain) and 1's (brain). The brain masks for PET volumes were defined by all voxels above a user-selected threshold; the mask boundaries defined the brain surface. The PET threshold was initially set at 45% of the maximum voxel value and was then interactively adjusted by the user to avoid removing white matter and low activity regions within the brain. The coordinates of the surface points in each volume were then converted to centimeters to adjust for different voxel sizes. For MRI-PET registration, MRI and PET volumes were designated as "head" and "hat," respectively.

The surface-fitting routine developed by Pelizzari et al. (7) was used to transform (translate, rotate, scale) points on the "hat" surface so that they "best" fitted the "head" surface. The surfaces were fitted by iteratively minimizing a residual equal to the mean squared distance between transformed "hat" points and the "head" surface. The iterative minimization algorithm was typically run 6–10 times; registration parameters from the previous run were used as starting parameters for the next run until a stable minimum was found. We found that this process almost always resulted in smaller residual fit errors than a single application of the algorithm. Upon completion of the surface fitting, the HH-derived scaling factors were dis-

carded, and new scaling factors were calculated using the voxel size for each volume.

Equivalent Internal Points (EIP)

Twenty-four easily identifiable points were defined in both MRI and PET images: inferior surface of anterior temporal lobe (L/R; left/right hemisphere locations, respectively), temporal pole (L/R), posterior surface of cerebellar hemisphere (L/R), occipital pole (L/R), most anterior of L/R frontal poles, center of head of caudate nucleus (L/R), center of putamen (L/R), center of thalamus (L/R), lateral surface of cerebellar hemisphere (L/R), and most lateral point of parietotemporal cortex (L/R). The remaining five points are all located on the midsagittal plane: the superior border of corpus callosum at the level of the anterior thalamus, genu, splenium, center of the fourth ventricle, and center of pons. Axial, coronal, and sagittal slices from the scans to be registered were displayed using our standard display software (25), and the equivalent points were interactively marked with a mouse-driven cursor. The two sets of equivalent points were translated to match their centers of mass (centroids) at the origin. Procruste's algorithm (8) was then used to compute the rotation matrix that minimized the root mean square error between the two sets of translated equivalent points.

Ratio Image Uniformity (RIU)

For the RIU technique, the translation and rotation parameters were determined such that they minimized the voxel-by-voxel variance of a "ratio image." The "ratio images" were created using the techniques described by Woods et al. (10) for MRI-MRI and PET-PET and by Woods et al. (11) for MRI-PET. The MRI volumes were "stripped" prior to calculating a "ratio image." We used the RIU algorithms and programs described by Woods (12) to iteratively minimize the ratio image variance using a modified Newton–Raphson method based on explicitly computed derivatives of the variance with respect to each of the registration parameters.

RESULTS

Headholder MRI and Transformed MRI Scans

The mean (±SD) and maximum VREs plotted in Fig. 1 represent the best registration performance achieved by all techniques. The mean VRE is of the order of the transverse pixel size of 1 mm except for SZF and RIU, which have significantly smaller VREs. The maximum errors approach the slice width of 3 mm. The greatest interobserver variation

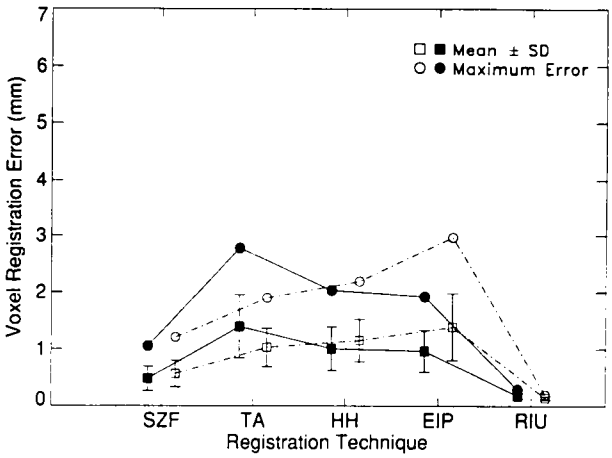

FIG. 1. MRI-MRI mean and maximum voxel registration errors (mm) for all brain voxels of six "headholder MRI scans" after registration with their "transformed MRI volumes" (mathematically misaligned from the "headholder MRI scans"). Results are plotted for two observers (open and filled symbols) and for each of five registration techniques: SZF, stereotaxic Z frames; TA, Talairach coordinates; HH, "head and hat"; EIP, equivalent internal points; RIU, ratio image uniformity.

was found for TA and EIP, the two techniques that are sensitive to user selection of neuroanatomical locations. The modified TA approach performed as well as HH and EIP for mean errors in this ideal MRI-MRI dataset (in contrast to the original TA technique reported in ref. 5), but the order of the maximum errors is reversed between techniques for the two observers. Slight interobserver differences in the mean RIU results probably reflect slight differences in the interactive creation of stripped MRI volumes.

Simulated PET and Transformed PET Scans

Figure 2 presents the VRE results for the simulated and transformed PET volumes that are characterized by reduced resolution and increased noise relative to the original MRI scans. There is no SZF result because the PET simulation did not include these external fiducial markers. For HH and EIP the mean VRE is slightly greater than in Fig. 1, and there is less interobserver variation. The poorest alignment is provided by TA, although the errors are still better, on average, than the voxel size of ~3 mm^3. Implementation of the indirect technique for finding the AC-PC line described by Friston et al. (23) gave poorer results than those plotted in Fig. 2. We found it necessary to use a weighted regression to reduce the impact of variability in the position of the occipital pole and adjust the average PET regression lines inferiorly (the four points being regressed were dropped inferiorly) with a slight posterior rotation to coincide with the AC-PC lines identified from the original MRI volumes.

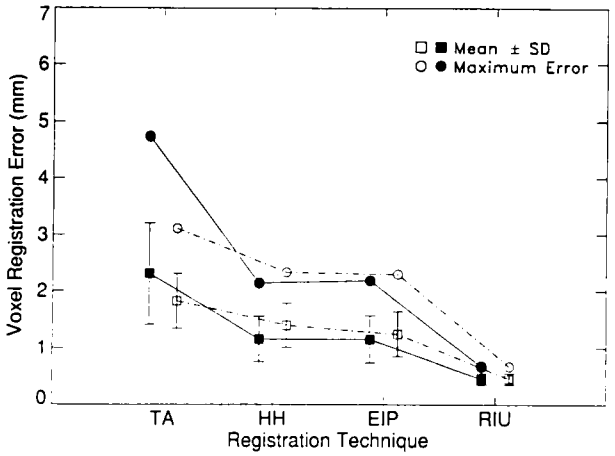

FIG. 2. PET-PET mean and maximum voxel registration errors (mm) for all brain voxels of six "simulated PET scans" after registration with their "transformed PET scans" (mathematically randomly misaligned from the "simulated PET scans"). Results are plotted for two observers (open and filled symbols) and for each of four registration techniques: TA, HH, EIP, RIU (see legend to Fig. 1).

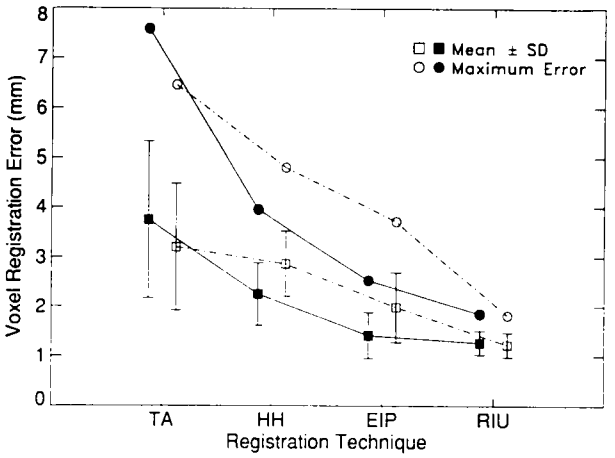

FIG. 3. MRI-PET mean and maximum voxel registration errors (mm) for all brain voxels of six "headholder MRI scans" after registration with their "transformed PET scans" (mathematically randomly misaligned from "simulated PET scans"). Results are plotted for two observers (open and filled symbols) and for each of four registration techniques: TA, HH, EIP, RIU (see legend to Fig. 1).

The RIU results represent registration of two sets of six simulated and six transformed PET-PET pairs; in addition to different noise distributions, one set had identical and the other different axial distributions of mean slice values to test the impact of changing spatial distributions across a PET-PET pair. The impact of these axially varying mean slice differences was found to be quite small, and the registration results from both sets were combined to produce a single group mean (±SD) and maximum VRE for each observer (Fig. 2). The RIU registration results for both sets of volume pairs were significantly better than for any other technique.

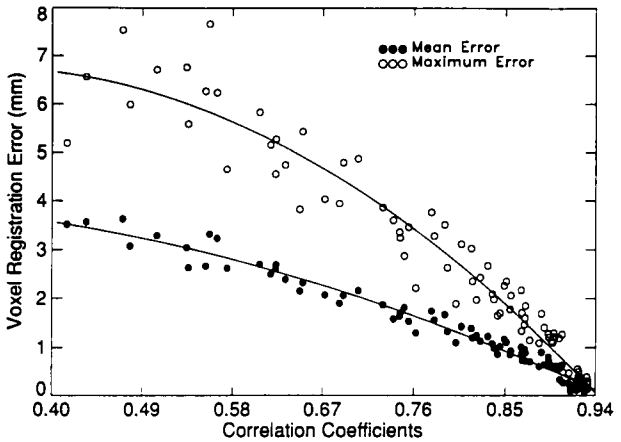

FIG. 4. Mean and maximum voxel registration errors (mm) for the MRI-MRI pairs of Fig. 1 plotted as a function of the correlation coefficients of paired volumes (intracranial voxels only) after registration. Mathematically misaligned pairs were created to increase the number of points with lower correlation coefficients. The solid lines represent quadratic regression fits to the two datasets.

Headholder MRI and Transformed PET Scans

For the MRI-PET registration results plotted in Fig. 3, the VREs of all techniques have increased compared with those of MRI-MRI (Fig. 1) and PET-PET (Fig. 2). The TA provides the worst registration with mean VREs slightly larger than the PET voxel size (~3 mm^3), and RIU has mean VREs only slightly larger than the transverse MRI pixel size (1 mm^2). Based on mean and maximum VREs, EIP is better than HH, and RIU is the best technique, particularly with respect to the maximum-to-mean VRE ratio. However, the advantage of RIU is more marked for one observer (open symbols) than the other (filled symbols). As might be expected, the maximum VRE is a more sensitive measure of improved registration performance across techniques than the mean VRE and improves faster from left to right in Fig. 3.

Headholder and Repositioned MRI Scans

Figure 4 presents the relationship between correlation coefficients and mean and maximum VREs, based on the correlation coefficients and VREs of headholder and transformed MRI pairs. Quadratic regression is seen to provide a reasonable fit to the data points. Figure 5 represents the VREs obtained from the correlations of the registered image volume pairs using the fitted curves in Fig. 4. When aligning MRI-MRI pairs, the real data results in Fig. 5 reflected the ideal simulated results in Fig. 1; i.e., there was little difference between TA, HH, and

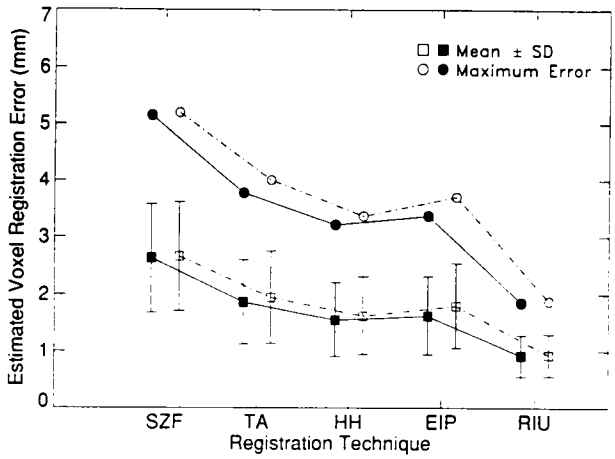

FIG. 5. Estimated MRI-MRI mean and maximum voxel registration errors (mm) for all brain voxels of six "headholder MRI scans" after registration with their "repositioned MRI scans." Results are plotted for two observers (open and filled symbols) and for each of five registration techniques: SZF, TA, HH, EIP, RIU (see legend to Fig. 1).

EIP. Even when a careful attempt was made to reposition subjects' heads within customized foam molds, the Z fiducial markers, which are fixed in relation to the molds, did not perform as well as any other registration technique. As for all previous results, the RIU technique was significantly better than all other techniques, particularly with respect to the maximum VREs.

DISCUSSION

Registration errors are accurately known for our transformed MRI and transformed PET volumes, and there are no error sources of significant size that have not been accounted for. Under these conditions, we are obtaining near optimal performance for the six techniques evaluated, and this performance is primarily dependent on the registration procedures, image resolution and noise, and the trilinear interpolation used to generate new volumes from known transformation matrices. The transformed MRI volumes are ideal in the sense that they simulate a perfectly stable MRI signal for MRI-MRI pairs. The simulated PET volumes are ideal in the sense that the PET signals are more highly correlated with the underlying anatomy (because they are directly derived from it) than would normally be expected in real FDG-PET scans.

The best overall registration results were obtained with MRI-MRI pairs (Fig. 1), which have the most accurately localized anatomical information because of their high spatial resolution. The similar performance of HH and EIP techniques with this idealized dataset is also found for PET-PET image pairs, but EIP has improved performance for MRI-PET registration (Fig. 3); however, the performance of SZF and TA, and RIU to a lesser extent, is dependent on the type of image pair being registered. Given the near ideal nature of our MRI-MRI datasets (an ideal dataset would have a resolution volume of <1 mm^3), the quality of the registration obtained by locating the midsagittal plane and the two commissures for TA is nearly as good as that obtained using 24 points for EIP or surface information for HH (Figs. 1 and 5).

As image volume resolution is decreased and noise increases in PET-PET pairs (Fig. 2) relative to MRI-MRI pairs (Fig. 1), the registration error increases for all techniques. Nevertheless, average errors are significantly smaller than the PET voxel size (~ 3 mm^3), and errors for HH and EIP increase by only a small amount compared with those illustrated in Fig. 1. This result is expected since the surfaces identified for HH are only minimally dependent on the underlying voxel size (19), and any tendency to randomly misplace individual anatomical locations in EIP is averaged out over 24 points (8). Talairach errors increase more than those associated with the other approaches since the PET-PET registrations rely on indirect estimates of the locations of the commissures. At the same time, RIU maintains average errors that are significantly smaller than those associated with the other techniques. The advantage of the RIU approach persists when each simulated PET volume of a PET-PET pair has a different spatial distribution of voxel values, e.g., the different axial distributions of slice means used in our simulations. This result indicates that RIU is relatively robust to functional differences between image volumes, as demonstrated by Woods et al. (10) for images with and without activation foci.

The increased registration errors for MRI-PET pairs (Fig. 3) compared with PET-PET pairs (Fig. 2) demonstrate that indirectly registering PET-PET pairs—first aligning them to MRI volumes and then aligning the MRI volumes—is less accurate than directly aligning the PET volumes. One source of this additional error was highlighted by our finding that better registration results were obtained for HH if the scaling factors were allowed to vary for PET-MRI registrations (19,20)—the scaling factors helped compensate for the difficulty in finding similar brain boundaries in anatomical–functional image pairs. Observers also reported that they found it easier to confidently identify equivalent points for the PET-PET pairs than for the PET-MRI pairs. The increase in TA errors for MRI-PET pairs (Fig. 3) compared with PET-PET pairs (Fig. 2) illustrates that observer error in reproducibly identifying the four points used to indirectly locate the AC-PC line in PET-PET pairs was smaller, on average, than the error between the estimated AC-PC line from PET and the "true" AC-PC line from MRI.

For MRI-PET pairs, RIU produced the best

alignment (Fig. 3), although it was no longer as clearly superior to EIP for both observers as it was for MRI-MRI (Fig. 1) and PET-PET (Fig. 2). Compared with other registration techniques, the trend for the RIU advantage to be successively reduced for MRI-MRI, PET-PET, and MRI-PET comparisons reflects the fact that if the image volumes being aligned have sufficiently different spatial distributions of voxel values (e.g., anatomical versus functional), minimizing the ratio image variance may cease to be a superior optimization strategy. The question remains: Are "significant" functional or functional–anatomical image differences rare or common? This question was not addressed by our datasets and should be further investigated.

The well defined relationship between mean and maximum voxel registration errors and correlation coefficients for headholder and transformed MRI volume pairs is illustrated in Fig. 4 and was used to estimate the registration errors plotted in Fig. 5. Compared with the errors illustrated in Fig. 1, the overall increase in average errors of 0.5–1.0 mm in Fig. 5 probably results from lower correlation coefficients due to signal variations in MRI volumes acquired at different times compared with the perfect signal stability between the headholder and transformed MRI volumes.

The time required to complete the alignment of image pairs listed in Table 1 ranged from a minimum of ~5 min for MRI-MRI with SZF and for PET-PET with RIU to a maximum of 30 min. The actual running of the registration algorithm followed by trilinear interpolation to form the registered volumes often took only a few minutes; it was the stripping of both volumes and the identification of the TA system that accounted for most of the time listed in the table. Although the RIU algorithm has been termed "automated" (11), we found that the time required to strip the skull and scalp from the MRI volumes (~10 min for 50 slices) was roughly equivalent to that required for an experienced user to identify the 24 EIP locations with well designed software (25).

TABLE 1. *Approximate times (min) required to register MRI-MRI, PET-PET, and MRI-PET volumes using five different registration techniques*

Technique	MRI-MRI	PET-PET	MRI-PET
SZF	5	—	—
TA[a]	20	20	20
HH[b]	30	10	20
EIP	30	30	30
RIU[b]	30	5	20

All registrations were performed using a package of IDL-based software tools designed to facilitate user interaction (22) and run on a SUN 4/75 (SPARC II) with 64 Mbyte of memory. For abbreviations see the text.

[a] Approximately 10 min is required to define Talairach coordinates for each volume.

[b] Ten minutes is required for each MRI volume to "strip" extrameningeal tissues and air.

Our experience using TA provides several useful insights into the errors involved in using Talairach coordinates. The TA errors for MRI-MRI pairs (Fig. 1) may be interpreted as a measure of the ability to reproduce Talairach coordinate locations in repeated MRI scans of the same subject. On average, we can expect identical coordinates within the two volumes to be within ~1 mm of each other. This 1 mm error resulting from the two stage fitting procedure used represents a reduction of 50% in the mean VRE (≈2 mm) that we previously reported for direct implementation of the Friston et al. technique (23). Our indirect estimate of the AC-PC line (before the second stage correction to match the directly measured, average MRI AC-PC line) has the same features reported by others (24): The midpoint of the estimated AC-PC line lies posterior and superior to the true AC-PC midpoint with anterior convergence. Following estimation of the AC-PC lines in PET-PET pairs, Fig. 2 illustrates that the reproducibility of Talairach coordinate locations for repeated FDG-PET scans averages 2–2.5 mm. The TA errors in Fig. 3 are included only for consistency, as they represent the combined effect of identifying the AC-PC line in both scans of an MRI-PET pair; the normal procedure would be to directly identify the AC-PC line in the MRI scan and align the MRI-PET pair by a more accurate technique (24).

While a wide range of headholding techniques and various evaluation criteria have been proposed (26,27), our results (SZF in Figs. 1 and 5) and the experience of others (28) suggest that it is not possible to fix the head accurately enough relative to external fiducial markers to obtain registration results as good as those obtained by nonfiducial-based techniques. Despite the refinement of our measurement technique, which resulted in smaller SZF errors than were reported in ref. 5, the fiducial markers attached to a custom head mold remained the worst registration approach that we evaluated. Furthermore, the SZF results reported in Fig. 1 demonstrate that this is a positioning and not an algorithmic problem. Better results may be obtained when the fiducials can be rigidly fixed to the skull, but even then the brain may move relative to the skull between repeated scans (28). The potential for movement of the brain relative to the skull is the reason why we have used the brain surface rather than the inner table of the skull for registration with HH.

A number of registration error comparisons have been reported in the literature. Despite differences in measuring and calculating registration errors, these reports together with our results indicate that there are two general classes of techniques for registering intra- or intermodality, intrasubject image pairs. The first class includes HH, PA, and EIP.

Rusinek et al. (19) compared PA (6) and HH (7) as the most promising techniques for automated registration and concluded that PA and HH were both capable of achieving a registration accuracy of 1–2 mm. Pelizzari et al. (7) found errors due to HH ranging from 0.7 to 2.5 mm depending on the voxel sizes of the images being registered, and Turkington et al. (20) found that errors of ≤2 mm were readily obtained for most structures. In a comparison of HH with EIP, Pelizzari et al. (29) found landmark location errors consistent with registration errors of 1–2 mm. All these results are consistent with our own, indicating that the use of limited numbers of internal points (EIP) or surface (HH) and/or volume (PA) information may readily achieve mean registration errors of ≤2 mm with individual location errors on the order of 1–3 mm. Chamfer matching techniques (30) appear to generate errors of this order, as does matching three orthogonal views (31); these additional techniques belong to the same class as HH, PA, and EIP.

The second class of techniques utilizes a comparison of internal voxel distributions and includes the "stochastic sign change criterion" (SSC) (32,33), "sum of absolute differences" (SAD) (34,35), "ratio image uniformity" (RIU) (2,10–12), "mean square difference" (MSD) (32), and "cross-correlation coefficient" (CC) (36) cost functions. Published results and our experience indicate that techniques of this class can routinely produce registration errors as low as—or lower than—the registration errors associated with the first class of techniques. Hoh et al. (35) tested SSC and SAD and found that they achieved similar registration accuracy in both brain and body, while Eberl et al. (34) and Minoshima et al. (37) found SAD to have a clear advantage in the brain. The registration accuracy of RIU has been compared with that of SSC, SAD, MSD, and CC for F-DOPA PET images. Lin et al. (36) concluded that the MSD and RIU techniques performed best; however, Venot et al. (32) concluded that MSD and CC are sensitive to differences in the spatial distribution of voxel values. Recently, Collins et al. (38) have demonstrated that CC techniques may be quite stable when implemented as a multistep multiresolution scaling approach. In an earlier article, Collins et al. (39) demonstrated that PA followed by CC can achieve intrasubject misregistration errors across 48 landmark points of <0.2 mm. There is mounting evidence that some Class II registration techniques, which utilize optimization criteria based on matching summary statistics from internal voxel distributions, perform consistently better than Class I techniques that ignore such information. Published results and our experience indicate that SAD or RIU may be optimal for intrasubject brain registration tasks.

If registration to within a PET resolution volume (~5 mm^3) is all that is required for a particular imaging task, our results indicate that any of the techniques we evaluated will prove satisfactory. All but SZF and TA easily achieve registration errors within the limiting axial resolution (3 mm) of our MRI scans. The registration accuracy required by imaging tasks remains a subject of debate (40), although there is evidence that [^{15}O]water activation studies may benefit from Class II registration techniques that can achieve registration errors much smaller than a PET resolution volume (34,37). There appear to be no computational advantages for the Class I techniques (PA, HH, EIP) over RIU, which has a significant computational advantage for PET-PET registration (see Table 1). Given the subpixel errors achieved by RIU for MRI-MRI and PET-PET registrations, it seems unlikely that another Class II technique would produce better results for our datasets. However, SAD or RIU may produce the smallest subpixel errors for MRI-PET registrations and for PET-PET registration with very different image distributions.

Until we understand the tasks for which subresolution and/or subpixel registration errors are important and which Class II techniques produce the smallest errors, we plan to follow a two step procedure for intrasubject registration. We will identify the Talairach transformation for each image volume, and this transformation will then be used to initialize a RIU registration algorithm (12). Our recent experience (unpublished data) indicates that this Talairach initialization step may be necessary to obtain a useful RIU result when dealing with T1 MRI volumes and badly misaligned PET volumes. Identification of the Talairach coordinates as a preliminary step has four additional advantages: It may help to speed up RIU computation times; it provides a standard means of reporting anatomical locations; it allows for the calculation of the difference between TA- and RIU-derived transformation matrices as a quality control procedure to detect outlying registration errors; and Talairach coordinates alone may be adequate for some intrasubject registration tasks. We also anticipate the use of EIP when outlier registration errors are detected, when pairs of image volumes are abnormal or have sufficiently different distributions that we have no confidence in either Talairach coordinates or RIU optimization, or when "MRI stripping" is particularly difficult, e.g., some T1 MRIs.

Acknowledgment: We thank Charles Pelizzari, Ph.D., for providing the "head and hat" alignment software and Roger Woods, M.D., for providing the "ratio image uniformity" software. This work was supported by National Institutes of Health grants NS25563, DA07428, and NS25701.

REFERENCES

1. Vannier MW, Gayou DE. Automated registration of multimodality images. *Radiology* 1988;169:860–1.

2. Woods RP, Mazziotta JC, Cherry SR. Automated image registration. In: Uemura K, Lassen NA, Jones T, Kanno I, eds. *Quantification of brain function: tracer kinetics and image analysis in PET*. Amsterdam: Elsevier Science, 1993:391–400.
3. Fox PT, Perlmutter JS, Raichle ME. A stereotactic method of anatomical localization for positron emission tomography. *J Comput Assist Tomogr* 1985;9:141–53.
4. Fox PT, Mintun MA, Reiman EM, Raichle ME. Enhanced detection of focal brain responses using intersubject averaging and change-distribution analysis of subtracted PET image. *J Cereb Blood Flow Metab* 1988;8:642–53.
5. Anderson JR, Strother SC, Xu X-L, Bonar DC, Rottenberg DA. Error bounds for five registration techniques based on high resolution MRI. In: Uemura K, Lassen NA, Jones T, Kanno I, eds. *Quantification of brain function: tracer kinetics and image analysis in PET*. Amsterdam: Elsevier Science, 1993:429–38.
6. Alpert NM, Bradshaw JF, Kennedy D, Correia JA. The principal axes transformation—a method for image registration. *J Nucl Med* 1990;31:1717–22.
7. Pelizzari CA, Chen GTY, Spelbring DR, Weichselbaum RR, Chen CT. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *J Comput Assist Tomogr* 1989;13:20–6.
8. Evans AC, Marrett S, Collins L, Peters TM. Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems. In: Schneider RH, Dwyer SJ III, Jost RG, eds. *Proceedings of the SPIE: medical imaging III: image processing. SPIE* 1989;1092:264–74.
9. Talairach J, Tournoux P. *Co-planar stereotaxic atlas of the human brain*. New York: Thieme, 1988.
10. Woods RP, Cherry SR, Mazziotta JC. Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr* 1992;16:620–33.
11. Woods RP, Mazziotta JC, Cherry SR. MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 1993; 17:536–46.
12. Woods RP. *Automated image registration user's manual and program documentation*. Los Angeles: Neuropsychiatric Institute, UCLA School of Medicine, 1993.
13. Kearfott KJ, Rottenberg DA, Knowles RJR. A new headholder for PET, CT and NMR imaging. *J Comput Assist Tomogr* 1984;8:1217–20.
14. Peters TM, Clark JA, Olivier A, et al. Integrated stereotaxic imaging with CT, MR imaging and digital subtraction angiography. *Radiology* 1986;161:821–6.
15. Ballard DH, Brown CM. *Computer vision*. Englewood Cliffs: Prentice-Hall, 1982.
16. Ende G, Treuer H, Boesecke R. Optimization and evaluation of landmark-based image correlation. *Phys Med Biol* 1992;37:261–71.
17. Greitz T, Bohm C, Holte S, Eriksson L. A computerized brain atlas: construction, anatomical content, and some applications. *J Comput Assist Tomogr* 1991;15:26–38.
18. Thurfjell L, Bohm C, Greitz T, Eriksson L. Transformations and algorithms in a computerized brain atlas. *IEEE Trans Nucl Sci* 1993;40:1187–91.
19. Rusinek H, Tsui W-H, Levy A, Noz ME, Leon M. Principal axes and surface fitting methods for three-dimensional image registration. *J Nucl Med* 1993;34:2019–24.
20. Turkington TG, Jaszczak RJ, Pelizzari CA, et al. Accuracy of registration of PET, SPECT and MR images of a brain phantom. *J Nucl Med* 1993;34:1587–94.
21. Spinks TJ, Jones T, Bailey DL, et al. Physical performance of a positron tomograph for brain imaging with retractable septa. *Phys Med Biol* 1992;37:1637–55.
22. Bonar DC, Schaper KA, Anderson JR, Rottenberg DA, Strother SC. Graphical analysis of MR feature space for measurement of CSF, gray-matter, and white-matter volumes. *J Comput Assist Tomogr* 1993;17:461–70.
23. Friston KJ, Passingham RE, Nutt JG, Heather JD, Sawle GV, Frackowiak RSJ. Localisation in PET images: direct fitting of the intercommissural (AC-PC) line. *J Cereb Blood Flow Metab* 1989;9:690–5.
24. Evans AC, Marrett S, Neelin P, et al. Anatomical mapping of functional activation in stereotactic coordinate space. *Neuroimage* 1992;1:43–53.
25. Schaper KA, Anderson JR, Strother SC, Rehm K, Rottenberg DA. *Technical report: IDL software tools*. Minneapolis: Minneapolis VA Medical Center, 1993.
26. Strother SC, Perlmutter JS. Headholders for functional brain imaging. *J Cereb Blood Flow Metab* 1987;7:S16–8.
27. Bettinardi V, Scardaoni R, Gilardi MC, et al. Head holder for PET, CT, and MR studies. *J Comput Assist Tomogr* 1991;15:886–92.
28. Herholz K. Transcript of discussion following the digital atlas session. In: Uemura K, Lassen NA, Jones T, Kanno I, eds. *Quantification of brain function: tracer kinetics and image analysis in PET*. Amsterdam: Elsevier Science, 1993:436.
29. Pelizzari CA, Evans AC, Neelin P, Chen C-T, Marrett S. Comparison of two methods for 3D registration of PET and MRI images. *Proceedings of annual int. conference of the IEEE Engineering in Medicine and Medicine and Biology Society*, vol 1, 1991;221–3.
30. Magin J-F, Frouin V, Bendriem B. Nonsupervised 3D registration of PET and MRI data using chamfer matching. *Conference record of 1992 IEEE Nuclear Science Symposium and Medical Imaging Conference*, vol 2; 1993:1262–4.
31. Steinmetz H, Huang Y, Seitz RJ, et al. Individual integration of positron emission tomography and high-resolution magnetic resonance imaging. *J Cereb Blood Flow Metab* 1992; 12:919–26.
32. Venot A, Liehn J-C, Lebruchec J-F, Roucayrol J-C. Automated comparison of scintigraphic images. *J Nucl Med* 1986;27;1337–42.
33. Minoshima S, Berger KL, Lee KS, Mintun MA. An automated method for rotational correction and centering of three-dimensional functional brain images. *J Nucl Med* 1992; 33:1579–85.
34. Eberl S, Kanno I, Fulton RR, et al. Automatic 3D spatial alignment for correcting inter-study patient motion in serial PET studies. In: Uemura K, Lassen NA, Jones T, Kanno I, eds. *Quantification of brain function: tracer kinetics and image analysis in PET*. Amsterdam: Elsevier Science, 1993: 419–28.
35. Hoh CK, Dahlbom M, Harris G, et al. Automated iterative three-dimensional registration of positron emission tomography images. *J Nucl Med* 1993;34:2009–18.
36. Lin K, Huang SC, Melega W, et al. Performance evaluation of different optimization methods for intra-subject registration of FDOPA PET images. *J Nucl Med* 1993;5:183P.
37. Minoshima S, Koeppe RA, Fessler JA, et al. Integrated and automated data analysis method for neuronal activation studies using ^{15}O-water PET. In: Uemura K, Lassen NA, Jones T, Kanno I, eds. *Quantification of brain function: tracer kinetics and image analysis in PET*. Amsterdam: Elsevier Science, 1993:409–17.
38. Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* 1994;18:192–205.
39. Collins DL, Peters TM, Dai W, Evans AC. Model based segmentation of individual brain structures from MRI data. Proc. SPIE: Visualization in biomedical computing, 1992: 10–23.
40. Strother SC. Chairman's remarks on the digital atlas session. In: Uemura K, Lassen NA, Jones T, Kanno I, eds. *Quantification of brain function: tracer kinetics and image analysis in PET*. Amsterdam: Elsevier Science, 1993:486–8.