We report a system for PET–MRI registration that is improved or optimized in several areas: (1) Automatic scalp/brain segmentation replaces manual drawing operations, (2) a new fast and accurate method of image registration, (3) visual assessment of registration quality is enhanced by composite imaging methods (i.e., fusion) and (4) the entire procedure is embedded in a commercially available scientific visualization package, thereby providing a consistent graphical user interface. The segmentation algorithm was tested on 17 MRI data sets and was successful in all cases. Accuracy of image registration was equal to that of the Woods algorithm, but 10 times faster for PET–PET and 4 times faster for PET–MRI. The image fusion method allows detection of misalignments on the order of 2–3 mm. These results demonstrate an integrated system for intermodality image registration, which is important because the procedure can be performed by technicians with no anatomic knowledge and reduces the required time from hours to about 15 min on a modern computer workstation.

INTRODUCTION

Over the past decade, many papers on image registration have appeared in the medical imaging literature, particularly on the coregistration of brain images. Recent reviews of this literature by van den Elsen et al. (1993) and Maurer and Fitzpatrick (1993) show the rapid evolution of this methodological research. Despite these advances, there is still room for improvement in several areas. For example, most registration techniques require, or at least benefit from, a preprocessing step that segments the structural MRI images, slice by slice, to remove extracranial structures, such as scalp, meninges, and skull. Although automated methods have been described (e.g., Ashkar and Modestino, 1978; Kass et al., 1988; Geman et al., 1990; and Davatzikos and Prince, 1995), this step is usually performed by manual drawing of boundaries. Manual segmentation requires drawing boundaries on up to 100 slices, a task that is tedious and which consumes several hours. Another area for potential improvement is the speed and convergence properties of the registration algorithms themselves. Still another problem is the evaluation of the accuracy of the registration. This evaluation should consider the quality of the image data and provide a quantitative estimate of the accuracy of registration. We are not aware of any analytic test which can provide such information on an individual subject's registration.

Image registration is used in critical applications, such as treatment planning. Unfortunately application of image registration techniques to routine clinical and research studies has been limited by practical issues such as those discussed above. In fact, because of cost considerations, it is our opinion that completely unsupervised methods will be necessary before image registration can reach its potential. The purpose of this paper is to report several developments which can help move the technology closer to this ideal. To that end, we present in this paper (1) a fast automatic method for scalp/brain segmentation, (2) an improved (faster) algorithm for image registration, and (3) an image fusion technique capable of detecting misregistrations of 2–3 mm magnitude. The techniques were validated/evaluated for PET–PET and PET–MRI registration. The advantages of these techniques are demonstrated by an integrated system (RS) recently developed in our laboratory.

METHODS

General

Software was written in ANSI C to extend the capabilities of a commercial scientific visualization package, AVS (Advanced Visual Systems, Inc., Waltham, MA). An integrated system, which we refer to as RS to denote registration software, was constructed using the network editor and user interface facilities of AVS.
The module library of AVS was extended to include 3D erosion, dilation, and connected components analysis as well as a number of modules designed to provide process synchronization and control logic. RS supports the following applications: (1) image readers with optional data cropping for PET and MRI, (2) brain surface segmentation, (3) image registration, PET-PET and PET-MRI, and (4) fusion.

Segmentation of Scalp/Brain Surface

The image volume is resampled to cubic voxels by trilinear interpolation. The volume histogram of gray levels is computed and used to manually select upper and lower intensity thresholds, removing low level noise outside the brain and bright extracranial voxels. After the thresholding operation “bridges” may remain, connecting the extracranial tissues to the brain surface. Binary erosion, based on a 3D structuring element, can break these bridges separating the image volume into clusters of voxels representing brain and extracranial tissues (Morris et al., 1993). Connected components analysis is employed to find the largest cluster, i.e., the brain. Binary dilation, performed on the brain cluster, restores the brain surface. It should be noted that the thresholding operation may remove voxels from the brain cluster, leaving holes that are only partially filled upon dilation. Indeed, missing voxels have been observed, particularly in the ventricular spaces and other fluid-filled spaces. However, the number of voxels involved is negligible and the holes are ignored in the registration process.

The two-dimensional connected components algorithm of Haralick (1981) was extended to three dimensions by altering the kernel and applying it in top-down and bottom-up passes in x, y, and z. The structuring element for binary erosion and dilation was a three-dimensional cross. Erosion and dilation were accelerated by implementation at the bit level using volume-shifting operations and logical operators as described by Haralick and Shapiro (1992) rather than a voxel-by-voxel application of the erosion/dilation kernel through the volume. The computational elements (two passes of erosion and dilation and connected component analysis) of this procedure on volumes of 256 × 256 × 180 voxels require less than 15 s on a 100-MHz workstation. Accordingly, adjustment of the brain/scalp segmentation parameters (e.g., the number of erosion/dilation passes and/or the kernel size) can be based on visualization of the resulting segmentation and thus allows fine control of the segmentation process. However, it should be noted, for a given pulse sequence and imaging protocol, all erosion/dilation parameters can be fixed; only the lower threshold need be adjusted.

Image Data

PET scans were acquired with a 15-slice GE-4096+ PET camera. A complete description of the imaging properties of this instrument has been given by Rotkops et al. (1990). Data were corrected for uniformity, photon attenuation, random coincidences, and dead time. Transverse section images were reconstructed with a conventional filtered back-projection algorithm to yield a final in-plane resolution of 8 mm FWHM. Both FDG and rCBF data were used in these studies. In order to improve the axial sampling of FDG, data were acquired in interleaved scan pairs by indexing the scanner bed 29.25 mm. This procedure resulted in a reconstructed image set with dimensions 128 × 128 × 30. Single scans were used with the rCBF data, resulting in image volumes with dimensions 128 × 128 × 15.

MRI data were acquired with GE-SIGNA scanners as 3D-SPGR volumes (Tr = 25 ms and TE = min) with slice thicknesses of 1.5, as part of the subjects routine diagnostic work up. Coronal data sets of 256 × 256 × 124 were reoriented to the transverse projection before use in the registration software.

A total of 23 data sets were available for evaluating PET–PET registration and 17 pairs, PET and MRI, were available for evaluating PET–MRI registration.

Image Registration Methods

The Woods method (Woods et al., 1992) for PET–PET image registration is based on the concept that two image volumes are in registration when the variance of the pixel-by-pixel ratio is a minimum. This assumes the spatial intensity patterns are similar for the two volumes, which is well justified for intramodality registration, e.g., PET to PET or MRI to MRI. When registering functional images, such as blood flow or glucose metabolism (i.e., PET), to images with T1- or T2-weighted volumes, this assumption is only approximately true, and the concept of minimum variance of the pixel-by-pixel ratio was extended to minimize a weighted sum of standard deviations of the PET voxels within the partitions defined by a segmentation of the MRI intensities into 256 levels (Woods et al., 1993).

We have developed a new method, based on the principles described by Woods and colleagues, the goal of which was to reduce the time required to estimate the registration transformation parameters. We denote the intensity values of the reference volume by \( S \), the intensity values of the image to be registered as \( I \), and the voxel coordinate by \( r \). For PET–PET registration, we minimize the cost function

\[
SSQ_1 = \sum_r \left[ S(r) - \frac{\bar{S}}{I} \cdot I(r) \right]^2,
\]

where \( \bar{S} \) and \( \bar{I} \) are the mean values of \( S \) and \( I \), respectively.

For PET–MRI registration we use a more complex cost function, involving partition of the reference vol-
volume. Let $r_p$ represent all voxels in partition $p$, $I(r_p)$ represent the intensities in partition $p$ and $I_p$ represent the mean value of the voxel intensities in partition $p$. Then the PET–MRI registration is computed by minimizing the sum of squares with respect to rotation and translation.

$$SSQ_2 = \sum_p w_p^2 \sum_{r_p} \left(1 - \frac{I(r_p)}{I_p}\right)^2,$$

where $w_p$ is a weighting function chosen to normalize the residuals. If we choose $w_p^2 = n_p^2/(n_p - 1)(\sum_p n_p)^2$ with $n_p$ denoting the number of voxels in partition $p$, then $SSQ_2$ is closely related to the cost function used by Woods et al. However, $SSQ_2$ cannot be considered as a sum of within partition variances with $p - 6$ degrees of freedom for the purpose of least squares estimation because the residuals are not properly defined. Instead, $SSQ_2$ is viewed as a sum of residuals with $\sum_p n_p - 6$ degrees of freedom.

An AVS module was developed to apply the theory described above to PET–PET and PET–MRI registration. Least squares fitting was done with the Levenberg-Marquardt method, based on the code described by Press et al. (1989), estimating six parameters: pitch, roll, yaw, $dX$, $dY$, and $dZ$. Absolute scale, i.e., voxel size, was assumed to be known. All parameter values were initialized to zero. Then, as described previously by Woods et al. (1992), a stepwise volume subsampling strategy was used to refine the fitting: Fits were performed using every $81^{st}$ voxel. After convergence at this level, the results were used to initialize the next sampling level, every $27^{st}$ voxel, and so on, until on the last step all voxels were included in the calculations. After registration, data were resliced using the fitted parameters and trilinear interpolation.

Evaluation and validation of the RS image registration modules were done by comparison with the AIR package, a widely used system, which has previously been validated by Woods et al. (1992, 1993) and Strother et al. (1994). Both RS and AIR were run on 23 volume pairs of FDG studies to evaluate PET–PET registration and on 17 volume rCBF–SPGR pairs to evaluate PET–MRI registration.

Registration performance was measured in several ways. We compared the fitting parameters obtained by RS with those of AIR, determining the mean, maximum, and rms differences over the ensemble of runs. We also, compared the convergence properties of the two methods to determine if RS could achieve results comparable to AIR in less time.

Visualization Methods

Registration algorithms have been validated using simulated data, phantoms or fiducial markers, but no accepted method for judging the accuracy of an individual registration has been described. To provide at least a gross evaluation of registration accuracy we have developed fusion methods using standard modules from the AVS library along with custom extensions which synchronize the independent processes. The fusion uses the AVS “composite” module which computes the fusion color table from a combination of the foreground and background image RGB color separations according to the formula

$$\begin{align*}
FUSION RED &= \text{foreground(red)} \times \text{alpha} + \\
&\quad \text{background(red)} \times (1 - \text{alpha}) \\
FUSION GREEN &= \text{foreground(green)} \times \text{alpha} + \\
&\quad \text{background(green)} \times (1 - \text{alpha}) \\
FUSION BLUE &= \text{foreground(blue)} \times \text{alpha} + \\
&\quad \text{background(blue)} \times (1 - \text{alpha}) \\
LUMINANCE &= 0.299 \times \text{red} + 0.587 \times \text{green} + \\
&\quad 0.114 \times \text{blue}.
\end{align*}$$

For example, in fusing PET and MRI, we calculate the luminance of the MRI image and use it to control the proportion (alpha) of foreground/background contrast. In this 2D fusion mode, three views are provided, transverse, coronal and sagittal. Widgets control slice selection, image contrast, color mapping, and the proportion of PET and MRI in the fusion. Alternatively, the operator can select three orthogonal views passing through a chosen point with the computer mouse.

We used simulation in order to gauge the effectiveness of these visualization methods for assessing accuracy. Known translation was applied in the axial, superior–inferior, and left–right directions in increments of 0 or 2 mm, resulting in translations as large as 2.8 mm. Three readers, who were not informed of the goals of the project, were asked to detect whether the images were in registration and to give the direction the misalignment. Readers were asked to detect errors by roaming through sets of transverse, coronal, and sagittal fusion images.

RESULTS

Segmentation of Scalp/Brain Surface

Figure 1 shows a montage of segmented MRI volumes, each presented as a volume rendered ray-traced volume, to illustrate the quality of the process. No manual editing was employed.

Image Registration

Results from RS and AIR for PET–PET registration were compared for subsampled volumes, using 1/81, 1/27, 1/9, 1/3, and 1/1 voxels. Increasing the sampling beyond 1/27 points did not effect the registration parameters significantly for either AIR or RS. Table 1 shows the mean and RMS difference between AIR and RS with 1/27 point sampling. Figure 2 graphically depicts the comparison of total displacement (defined as $\sqrt{dX^2 + dY^2 + dZ^2}$) for the two methods. Figure 3
shows the comparison of pitch angle for RS versus AIR. The maximum difference in displacement for the two methods was 1.1 mm, and the maximum difference in pitch angle was 1.1°. Although the registration parameters agree closely, RS was on average 10 times faster than was AIR.

Results for PET–MRI registration from RS and AIR were examined for subsampled volumes, using every 1/81, 1/27, 1/9, 1/3, and 1/1 voxels. Transformation parameters, three angles and three displacements, converged for both methods for 27-point sampling; increasing the sampling led to negligible changes in registration parameters. The results are summarized in Table 2. For AIR the maximum difference in displacement for 81- and 27-point sampling was 9.64 mm, while the corresponding maximum for RS was much smaller, 0.50 mm, indicating better convergence properties for the least squares method. Similar improvements in convergence properties were observed for rotation angle: the maximum change in angle when increasing sampling from 81 to 27 was 14.71° for AIR and 0.53° for RS. After 1/27 voxel subsampling these data show excellent agreement between AIR and RS; however, RS was on average 4 times faster than AIR.

![Figure 1](image.png)

**FIG. 1.** Montage of nine brains, each segmented to remove extracranial structures. Brain images are displayed as ray-traced volume renderings. The segmentation process requires no manual editing.

<table>
<thead>
<tr>
<th>Sampling (1/voxel)</th>
<th>Pitch (deg.)</th>
<th>Roll (deg.)</th>
<th>Yaw (deg.)</th>
<th>dX (mm)</th>
<th>dY (mm)</th>
<th>dZ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (RS-AIR)</td>
<td>27</td>
<td>0.15</td>
<td>0.04</td>
<td>0.01</td>
<td>0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>RMS (RS-AIR)</td>
<td>27</td>
<td>0.76</td>
<td>0.21</td>
<td>0.32</td>
<td>0.04</td>
<td>0.08</td>
</tr>
</tbody>
</table>

**TABLE 1**

PET—PET Registration, RS vs. AIR
FIG. 2. Comparison of total displacement (indicated by asterisk symbols) computed by two registration algorithms for PET–PET registration. The distance of points from the solid line indicates the deviation from perfect agreement between the two methods.

FIG. 3. Comparison of pitch angle (indicated by the asterisk symbols) computed by two registration algorithms for PET–PET registration. The distance of points from the solid line indicates the deviation from perfect agreement between the two methods.
Figure 4 compares the results of AIR and RS for the determination of the translation parameters. A plot of the total displacement distance obtained when the voxel sampling was set to 27 is shown. For the 17 image sets studied, the mean difference in displacement for AIR and RS was 0.28 mm and the rms difference between the two methods was 1.55 mm.

Figure 5 compares the results of AIR and RS for the determination of pitch angle. For the 17 image volumes studied, the mean difference in pitch angle for AIR and RS was 0.094° and the rms difference between the two methods was 0.32°.

The mean improvement in speed for the RS method was 3.9X, while the minimum improvement was 2.2 x and the maximum improvement was about 8 x.

**Evaluation of Fusions Technique**

Figure 6 shows a 3 x 3 montage of fused sagittal images. The middle panel is "perfectly" registered, whereas the images in the other panel are misaligned. For example, the images in the corners are misaligned by 2-pixel translations in the z- and y-directions, while the others are displaced with a single 2-pixel translation, right, left, up, or down. All three readers correctly indicated which data sets were misaligned; their accuracy in determining the direction of misalignment was 80%.

**DISCUSSION AND CONCLUSIONS**

Accurate location of the brain surface and/or removal of confounding extracranial structures (surface seg-
mentation) is often required in the processing or visualization of anatomic data. For example, removal of extracranial structures, such as scalp and meninges, is essential for computing a three-dimensional rendering of the brain surface and for accurate functioning of image registration algorithms (Pelizzari, 1989; Woods et al., 1993, Morris et al., 1993). With the evolution of volumetric MRI techniques, which produce 100 slices, or more, manual segmentation methods become the most important limitation to the wider use of coregistration methods in clinical and research applications.

The contrast between white and gray matter is important for our segmentation algorithm because it determines the shape of the intensity histogram, which, in turn, determines upper and lower level thresholds. Image contrast also effects the ability to do manual segmentation, in fact MRI gray levels do not always unambiguously identify intracranial and extracranial structures. Therefore, one must acknowledge that neither automatic nor manual methods can be perfect. Our segmentation method uses thresholding and morphological operations to automate the segmentation process. It is most sensitive to the low level threshold, whereas the setting of the upper threshold is not critical. Controlling the number of erosion/dilation iterations and the size and shape of the erosion/dilation kernel is essential for high quality segmentation. However, for a given pulse sequence and slice thickness all parameters except the low level threshold can be fixed.

One expects that registration algorithms based on information from all intracranial voxels will not be very sensitive to minor imperfections in the location of the brain surface.

The work on our registration algorithm was inspired by experience with the automatic image registration software developed by Woods and colleagues. Their algorithm was very reliable, but extended computation time was considered to be a limitation for some applications. In their algorithm the registration is determined by iteratively minimizing a cost function, one parameter at a time. Our first thought was that their code could be accelerated by using more efficient minimization techniques which adjusted all parameters simultaneously. However, when we substituted standard function minimizers, such as conjugate gradient or variable metric methods, in the AIR code, to our surprise, we found that they performed poorly and often would not converge to a solution. Graphical examination of the Woods cost function showed local minima associated with interpolation across voxel boundaries, which might explain some of the difficulties with the more classical minimization methods. Accordingly, we considered the possibility of least squares estimation with the hope that it might be less sensitive to the presence of these local minima while employing a search strategy that adjusts all parameters simulta-
neously. The least squares algorithm has proven to be as accurate as the Woods methods, but about 10 times as fast for PET–PET registration and 4 times as fast for PET–MRI registration. About a factor of 5 in speed PET–PET registration can be attributed to the least squares approach, since Woods et al. actually do their minimization twice, using each volume set alternately as the reference. The remaining speed improvement is attributable to the many fewer function evaluations required by the nonlinear least squares optimizer.

Studies on the accuracy of several registration algorithms show them to be accurate to 1–2 mm (Strother et al. 1994; Turkington, 1995). Image properties such as spatial and contrast resolution, statistical precision, geometric distortion, and spatial sampling all can be expected to affect the accuracy of the registration. As mentioned above, there is currently no theory which can take these properties into account to estimate the accuracy achieved in a given registration. Therefore, fusion techniques which can detect misalignment would be valuable for quality control or verification of the registration. Our method of fusion appears capable of detecting misalignments on the order of 2–3 mm.

With RS, the entire process of image registration is embedded in a commercially available scientific visualization system. The use of this system facilitates the development of portable computer code. The availability of high quality image display tools and a consistent graphical user interface facilitates the training of technicians and researchers in the use of the system. Most importantly, this integration of registration tools makes it possible to perform all phases of the process in about 15 min and about 1/3 of that time is used for handling the large image sets.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the Charles A. Dana Foundation Consortium on Memory Loss and Aging. Dr. Morri's contributions were supported, in part, by PHS Postdoctoral Training Grant 2T32CA09362.

REFERENCES


FIG. 6. A 3 × 3 montage of fused sagittal images. The middle panel is "perfectly aligned;" the other images are misaligned (see text for discussion). The bright red color near the brain surface indicates misalignment.
Parallel Computing Image Analysis (M. Onoe, K. Preston, Jr., and A. Rosenfeld, Eds.). Plenum, New York.


