A Rapid and Accurate Method to Realign PET Scans Utilizing Image Edge Information

Jesper L.R. Andersson

Uppsala University PET Centre, Department of Radiation Sciences and Subfemtomole Biorecognition Project, Uppsala University, Uppsala, Sweden

Movement during or between PET examinations is a common and serious problem. Consequently, there is a great need for rapid, accurate and robust methods to realign image sets. **Methods:** Derivative information from the image sets was used to extract areas containing edge information. Image similarity between a reference dataset and a misaligned dataset was evaluated for these areas. Powell's method for function minimization was used to find the set of translations and rotations along and around the axes that maximized image similarity. The method was validated by realigning image sets with a known misalignment. Image sets used for validation included brain studies using several different tracers and heart studies using labeled acetate or water. **Results:** The method was capable of realigning brain datasets using the same tracer with an accuracy of 0.2 mm and 0.2 degrees along and around all axes. The same accuracy was obtained for datasets with as few as a total of 800,000 counts. Brain studies utilizing different tracers with markedly dissimilar regional uptake patterns were realigned with an accuracy of 1.5 mm and 1.5°. Heart studies using water or acetate were realigned with an accuracy of 0.2 mm and 0.4° along and around all axes. Realignment of a heart study containing a large focal uptake defect against a dataset without defect produced errors no greater than 1.0 mm and 1.0°. **Conclusion:** The use of derivative information provides a useful method to accurately realign PET image sets. It is rapid and noise-insensitive enough to allow for its routine use in dynamic studies.

Key Words: positron emission tomography; image registration

J Nucl Med 1995; 36:657–669

Movement during or between PET examinations is a common problem that may seriously affect the analysis and interpretation of the examination. Several reports have been published lately describing how to register multiple PET datasets to each other (*1–5*). They all use the same basic methodology where one image set is resliced and compared to the other according to some image similarity criterion until the set of translations and rotations producing the best correspondence between the two image sets is found. Results from these studies all indicate that this is a feasible method with registration errors comparable to what could be expected using a head fixation system (*6*) or fiducial markers (*7*). Published methods differ in what criterion to use for image correspondence. Mintun et al. (*1*) reported good results using the stochastic sign change (SSC) (*8*) criterion. Eberl et al. (*5*) compared the SSC criterion with the sum of absolute differences (SAD) and found the latter measure to be superior, while Hoh et al. (*3*), on the other hand, used the very similar deterministic sign change (DSC) criterion (*8*) and the SAD in a comparative study, reporting essentially equal results for the two methods. Bacharach et al. (*4*) used the image correlation coefficient on transmission datasets with good results despite earlier reports that this should be inferior to the SSC criterion (*8*), and Woods et al. (*2*) finally used the standard deviation of a volume obtained by dividing one of the datasets with the other as their measure.

Important traits for an ideal method would be ease of implementation, speed of computation and a low sensitivity to noise. In this article, a method is presented that measures image similarity in which image volume is differentiated to yield information on where edges in the data are situated. The analysis then proceeds by evaluating the image correlation for voxels with a high derivative only. The important difference in comparison to the methods described above is the ability to restrict analysis to the edges in the images, which saves computational time and decreases noise sensitivity; this difference is essentially independent of similarity criterion. The technique represents a hybrid between the image similarity methods referred to above and surface fit techniques (*9*).

The method was investigated and optimized with respect to speed, registration accuracy and noise sensitivity. It was also applied to the realignment of cerebral blood flow (CBF) studies to ^{15}O carbon dioxide studies, an area where proper alignment is of great importance and misalignment may completely invalidate the data (*10*). Finally, data obtained for different tracers with markedly dissimilar regional uptake was realigned, a

Received Mar. 14, 1994; revision accepted Sept. 20, 1994
For correspondence or reprints contact: Jesper Andersson, MSc, Uppsala University PET Center, S-751 85 Uppsala, Sweden.

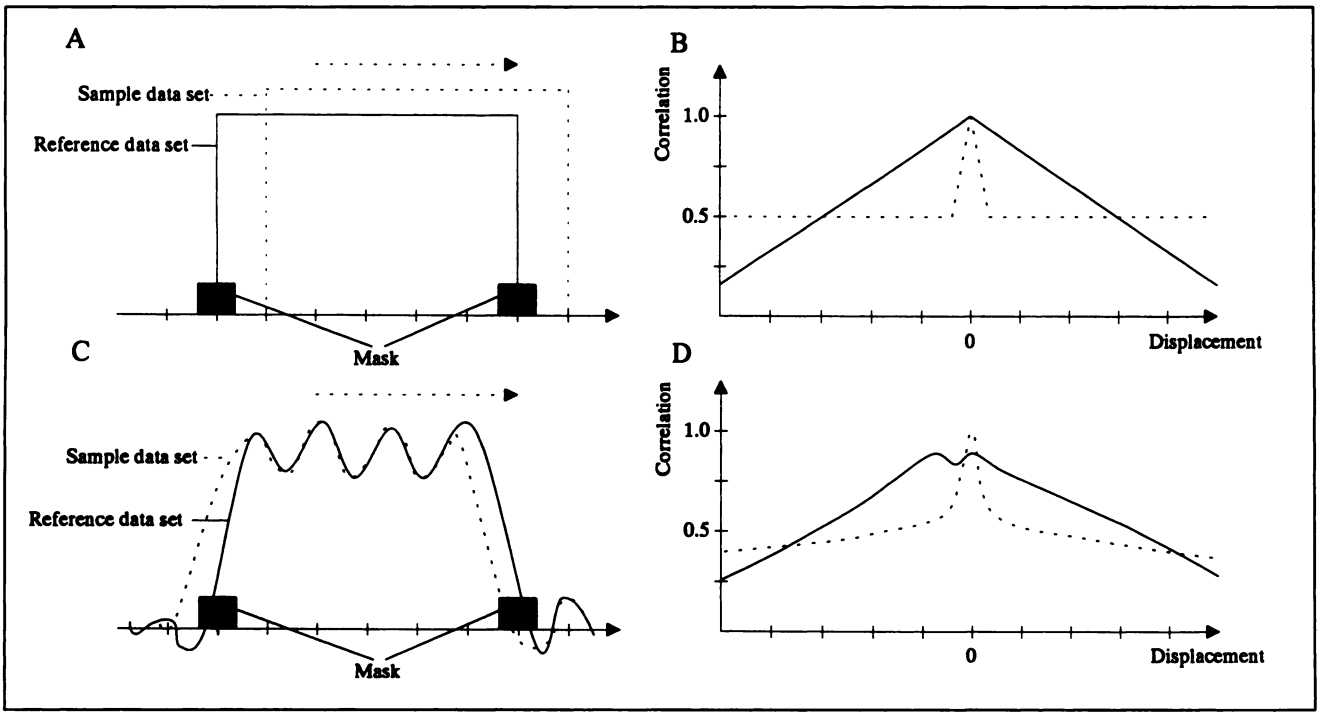

FIGURE 1. Schematic illustration of the effect of applying a mask around the edges of the reference dataset when evaluating its correlation with the sample dataset versus relative displacement between the datasets. (A) Two one-dimensional noise-free datasets where the sample dataset is translated relative to the reference dataset. (B) The resulting correlation coefficient, with and without mask, versus relative translation of the datasets in A. (C) Two one-dimensional datasets with noise added where the sample dataset is translated relative to the reference dataset. (D) The resulting correlation coefficient, with and without mask, versus relative translation of the datasets in C.

subject that has not been addressed and is of importance in multitracer protocols.

METHODS

Theory

The image correlation coefficient is a well-known measure of image similarity and has been used successfully to register PET transmission datasets (*4*). We suggest using the image correlation coefficient, but using it only on voxels having a high signal-to-noise ratio with respect to image orientation, that is voxels around edges in the images. This serves the twofold purpose of eliminating a large portion of the voxels in the image volume, thus saving computation time, and reducing the noise sensitivity of the method to improve its accuracy for low count studies. The rationale for this is described schematically in Figure 1. Figure 1A shows two one-dimensional, noise-free datasets of which the sample dataset is displaced relative to the reference dataset and a mask that has a nonzero value around the edges in the reference dataset. Figure 1B shows how the correlation coefficient varies as the sample dataset is moved relative to the reference dataset. The solid line shows the correlation evaluated for the whole datasets and the dotted line evaluated only for the parts with a nonzero mask. This indicates how the correlation coefficient varies faster close to the true solution when using the mask. In other words, when not using the mask, the evaluation of the correlation coefficient is based partly on many redundant values which do not contribute to the solution. Since the search for a solution normally takes us to within a few millimeters of the truth in the first couple of steps, it is not necessary to include voxels from more than a centimeter around edges in the images. Figure 1C shows the same two datasets with noise added. We added noise so that it happens to correlate well when the sample dataset is translated a given distance relative to the reference dataset. Figure 1D shows how the correlation coefficient varies as the sample dataset is translated relative to the reference dataset. Note how the solid line, representing the correlation evaluated for the whole-dataset, exhibits two separate peaks, one at the true position and the other at the position where the noise in the two sets was correlated. Figures 1C and 1D demonstrate the mechanism behind the noise sensitivity of these type of methods, regardless of which image similarity measure is used and thus applies equally well to the SAD criterion. Note also how the dotted line, representing the correlation evaluated only for the parts where the mask has a nonzero value, exhibits a single peak at the correct position.

The principle outlined above is incorporated into this method by applying a differentiating filter in the x, y and z directions to the reference dataset, producing three volumes with x, y and z derivatives of the images, respectively. These vector volumes are summed to yield a total derivative volume. A mask volume is created by thresholding the derivative volume so that five percent of the pixels in the derivative volume is above this threshold, obtaining a mask with the value one where the derivative is above this threshold and zero elsewhere. The whole method is shown in Figure 2 and summarized below:

1. Read the two image sets and denote one of them reference volume and the other sample volume.
2. Filter both datasets with a Gaussian filter.
3. Reslice the reference volume with a changed axial sampling.
4. Differentiate the reference datasets in x, y and z direction by applying a differentiating filter in each direction, sum these vector fields yielding a derivative volume.
5. Threshold the derivative volume to obtain a mask volume.

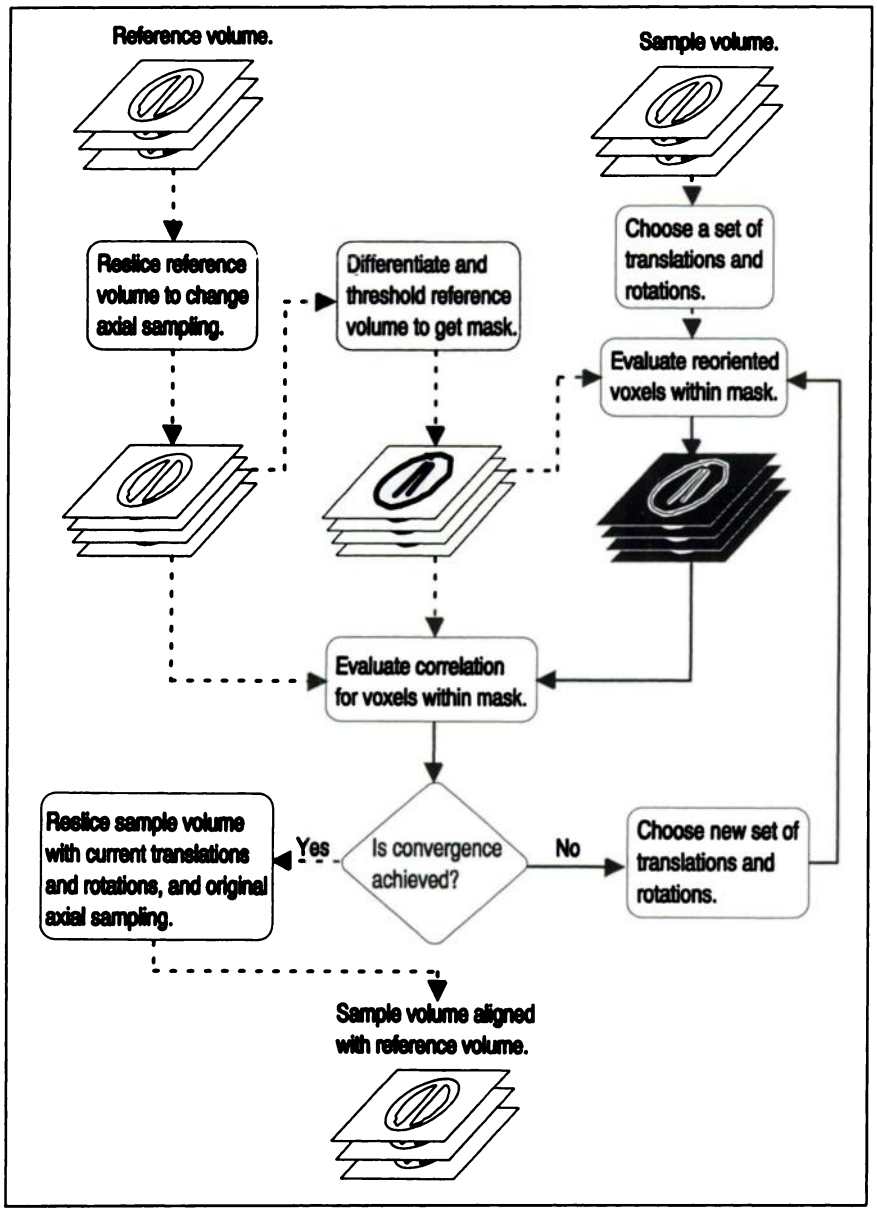

FIGURE 2. Description of procedure used to realign PET datasets. Dotted lines indicate operations that are executed once only and solid lines indicate operations that are executed once every iteration. Typically, the procedure goes through the loop indicated by the solid lines 200–400 times for every registration.

6. Choose a set of three translational and three rotational parameters.
7. Reslice the sample set, using trilinear interpolation, into those voxels of the reference set that are within the mask.
8. Evaluate the correlation coefficient for the voxels within the mask.
9. If convergence has occurred, reslice entire sample set into reference volume. Otherwise go to Step 6.

The minimization algorithm actually used is Powell's method (*11*) as implemented by Press et al. (*12*) with a few minor modifications to increase efficiency for this specific problem. If the reference volume is not resliced to yield a new axial sampling density as indicated in step three, there will be a tendency to obtain erroneous values for translation in the z direction. The mechanism behind this will be explained in the Discussion section.

When evaluating the correlation for the pixels within the mask further subsampling may be done so that not every pixel is used. This is done so that when every other pixel is used, every other x coordinate is used only. When every fourth pixel is used, every other x and y coordinate is used. When every eighth pixel is used, every fourth x coordinate and every other y coordinate is used and so on for further subsamplings. A more detailed description of the method can be found in the Appendix.

Experimental Methods

General Considerations. In order to estimate the performance of the method, datasets were needed where the subjects could be assumed to have been still during the course of the examination. Data were, in some instances, selected from examinations in which the subjects were sedated and on respirator ventilation. The cases where this was considered most critical were for the whole-body cases and for the cases where images obtained with different tracers were registered to each other. For the whole-body cases, there was no adequate fixation, and for data obtained with different tracers, there was a relatively long time span between acquisitions. When realigning brain studies obtained with

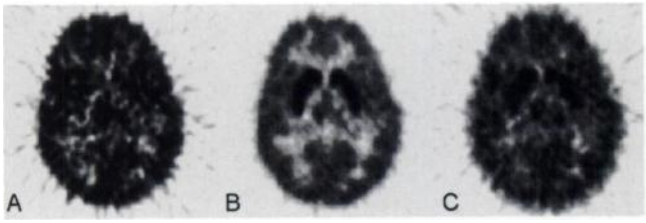

FIGURE 3. A central slice through the brain of the individual that was used for the multiple tracer experiment. It should be noted that, although not evident from these images, there is very little uptake of ^{11}C-NMSP in the cerebellum. (A) ^{15}O-water. (B) ^{11}C-NMSP. (C) ^{11}C-DOPA.

the same tracer, correctly aligned data were instead created by subdividing the frames of a dynamic study as described below.

Scanners. Brain studies were performed using a GE PC2048B (GE Medical Systems, Milwaukee, WI) scanner (*13*) and heart studies were performed using a GE PC4096WB scanner (GE Medical Systems, Milwaukee, WI) (*14*).

Computers. A VAX-station 4000/60 (Digital Equipment Corp., Maynard, MA) was used for all calculations presented in this article. The programs to perform the calculations were written in C and were linked to libraries written in FORTRAN delivered with the scanner and to routines copied from Press et al. (*12*).

Cerebral Blood Flow Studies. Approximately 15 MBq/kg body weight ^{15}O-water were injected intravenously, after which scanning was begun. The scanning protocol consisted of 17 frames acquired over 5 sec followed by two 20-sec frames. In one case, scanning was continued for an additional six frames of 20 sec each. Images were reconstructed with a 6-mm Hanning filter, 2-mm pixel size, and a transmission scan was used for attenuation correction.

Brain Oxygen Studies. The subject was unconscious and respirator ventilated. The respirator hose was connected to a gas inhalation system and ^{15}O-carbon dioxide was continuously fed to the subject. After an equilibration time, scanning was commenced and continued for five frames for 1 min. The examination was reconstructed with a 6-mm Hanning filter, 2 mm pixel size, and a transmission scan was used for attenuation correction. The subject also had a CBF study according to the protocol above.

Brain ^{11}C-NMSP and ^{11}C-L-DOPA Studies. One individual was sedated and 6 MBq/kg body weight of ^{11}C-L-DOPA were injected; images were then obtained. After adequate time for decay of the first tracer, a second study was performed using 7 MBq/kg body weight of ^{11}C-N-methyl-spiperone (NMSP). Data from 15 to 60 min after injection were summed for both tracers and reconstructed with a 4.2-mm Hanning filter, 2 mm pixel size; a transmission scan was used for attenuation correction. This subject also had a CBF study according to the protocol above with an injected activity of 10 MBq/kg body weight. Representative images for all three tracers are shown in Figure 3.

Heart ^{11}C-Acetate Studies. In an anesthetized and ventilated pig weighing 19 kg with chest dimensions of 25 × 25 cm and heart dimensions of 7 × 7 × 7 cm, 28 MBq/kg body weight of ^{11}C-acetate were injected and a dynamic scan consisting of 32 frames for 81 min was obtained. The study was reconstructed with a 4.2-mm Hanning filter, 4 mm pixel size, and a transmission scan was used for attenuation correction.

Heart ^{15}O-Water Studies. The same pig from the ^{11}C studies was used. Oxygen-15-water (45 MBq/kg body weight) was injected and a dynamic scan consisting of 29 frames for 5.5 min was obtained. The images were reconstructed as above.

Data Verification

To evaluate the method's ability to accurately reorient misaligned datasets, two presumably properly aligned datasets were used. One dataset was labeled reference and the software was reoriented by applying a random set of three translations in the range −5 to 5 mm and three rotations in the range −5 to 5 degrees. The other volume was labeled sample and was reoriented with our method to yield the best possible agreement with the reference volume. The set of translations and rotations estimated by our method was then compared to the known set. When repeated many times with new random reslicings, this procedure will test the existence of local maxima, but since it is based on two datasets from one individual only, its spread may not be representative of that encountered when applying the method to a large number of truly different datasets. The procedure described above will be referred to in the remaining text as synthetic reslicing.

To check the spread of the accuracy in realignment, multiple datasets, presumably aligned in pairs, obtained from different individuals were realigned with our method. Any deviation from zero in translation or rotation was assumed to be erroneous and the spread in error between the different individuals was used as a measure of the spread of the method. This procedure will be referred to as multiple realignment in the remaining text.

When realigning datasets obtained with the same tracer, presumably correctly aligned datasets were generated by dividing the frames of a dynamic study into two parts, one consisting of odd numbered frames and one of even numbered frames. These frames were then summed to yield two summation image files containing similar, but distinct data. This technique was not usable when realigning data obtained with different tracers.

The performance of the method was examined with respect to the following aspects.

The Effects of Initial Resampling of the Reference Image Set. Two datasets containing 5.6 million counts each were created from a dynamic CBF study on one healthy volunteer according to the method described above. To demonstrate the effects from no resampling of the reference dataset in the axial direction, the correlation between the two datasets was evaluated while moving the sample dataset in the axial direction relative to the other reference dataset. This was performed both with and without axial resampling of the reference image set. To demonstrate that the observed effects are not due to our choice of correlation as similarity criterion, the same was done using SAD as the similarity criterion.

The Effects of Prefiltering of the Image Sets. A CBF study was performed on one healthy volunteer. Scanning was continued until 245 sec after injection. Two pairs of properly aligned datasets were created from the dynamic study. One set with good statistics containing 5.6 million counts and one with poor statistics containing 800,000 counts. These datasets were synthetically resliced according to the procedure described above. This was repeated 50 times for each choice of filter FWHM.

Sampling Demands. Two datasets containing 5.6 million counts each were created from a dynamic CBF study on one healthy volunteer. These datasets were synthetically resliced according to the procedure described above. This was repeated 50 times for each choice of sampling density.

Noise Sensitivity. A CBF study was performed on one healthy volunteer. Scanning continued until 245 sec after injection. The data were rearranged into two parts, one part containing odd frames and the other containing even frames. Subsets of these parts were then summed in various combinations to yield paired

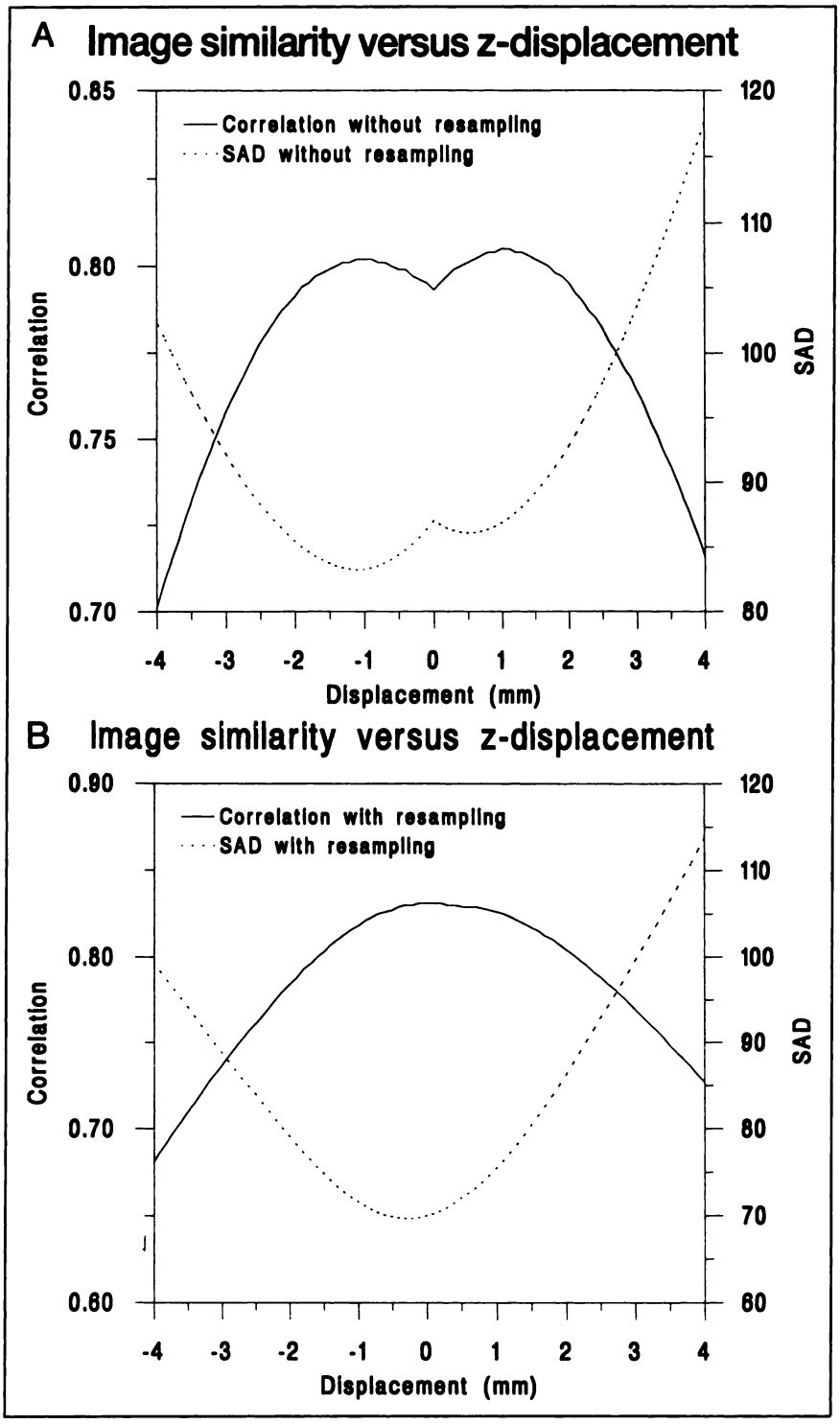

FIGURE 4. Correlation and SAD versus axial shift between reference and sample dataset. The data was obtained for two paired datasets containing a total of 800,000 counts each, and prefiltered with a Gaussian filter with FWHM 20 mm. Panel A shows the situation where the sample dataset is moved relative to the reference set and (B) shows what happens when the reference dataset is first resampled in the axial direction.

datasets, with a total number of counts for all slices ranging from 400,000 to 6 million. These pairs of datasets were synthetically resliced according to the procedure described above. This was repeated 50 times for each pair.

Registration Accuracy for CBF Studies. Seven healthy individuals had three CBF scans each according to the protocol above. From one of these individuals two datasets containing 5.6 million counts each were created according to the procedure described above. These were synthetically resliced 50 times with new random choices each time. From the remaining six individuals, two datasets containing 1.8 million counts each were selected from each scan, for a total of 18 pairs, and the best set of transformation parameters were found for each pair.

Registration Accuracy for CBF Studies with Local Perturbations. Two datasets containing 5.6 million counts each were created from a dynamic CBF study on one healthy volunteer. One of these datasets was manipulated to yield 20% activity increases in areas corresponding to what would be found in an activation study

TABLE 1
Registration Errors Versus Filter FWHM for High Count Studies

FWHM	Translation (mm)			Rotation (deg)		
	x	y	z	α	β	γ
0	0.13 ± 0.11	0.15 ± 0.10	0.17 ± 0.13	0.09 ± 0.09	0.16 ± 0.11	0.05 ± 0.06
10	0.11 ± 0.11	0.14 ± 0.11	0.12 ± 0.11	0.09 ± 0.08	0.13 ± 0.11	0.05 ± 0.06
20	0.14 ± 0.11	0.21 ± 0.10	0.10 ± 0.09	0.14 ± 0.11	0.13 ± 0.11	0.07 ± 0.05
30	0.14 ± 0.13	0.20 ± 0.15	0.09 ± 0.11	0.18 ± 0.13	0.10 ± 0.12	0.08 ± 0.10
40	0.26 ± 0.18	0.25 ± 0.15	0.09 ± 0.12	0.22 ± 0.14	0.17 ± 0.14	0.23 ± 0.18

with vibrotactile stimulation of the lip. This dataset was matched with the nonmanipulated dataset and synthetic reslicing was performed 50 times. Then one of the original two datasets was manipulated to simulate a massive middle cerebral artery infarction. All parts of the brain provided for by the left middle cerebral artery had its activity values reduced by 80%. This dataset was matched with the nonmanipulated dataset and synthetic reslicing was performed 50 times.

Registration Accuracy for CBF Studies to Oxygen Studies. Oxygen and CBF studies were performed in one individual according to the protocols above. A CBF dataset containing 4.2 million counts and an oxygen dataset containing 7.8 million counts was created. Synthetic reslicing was performed on the datasets 50 times.

Registration Accuracy for Brain Examinations with Markedly Different Tracers. One individual had CBF, NMSP and ^{11}C-DOPA studies according to the protocols above. These datasets were arranged into three groups of two. Synthetic reslicing was performed and repeated 50 times for each group.

Registration Accuracy for Myocardial Studies with ^{11}C-Acetate. Two groups of paired datasets were created from one dynamic study. One of the groups contained 4.4 million counts and the other 8.8 million counts. Synthetic reslicing was performed and repeated 50 times for each group.

Registration Accuracy for Myocardial Studies with Perfusion Defects. One pair of ^{11}C-acetate datasets, containing 4.4 million counts each, was created from one dynamic study. To mimic the situation from a perfusion reserve study, a large perfusion defect was simulated in one of the datasets by reducing the activity values by 90% in the lateral part of the left chamber wall. The volume of the defect was approximately 25% of the total left chamber wall. This dataset was matched with the nonmanipulated dataset and synthetic reslicing was performed 50 times.

Registration Accuracy for Myocardial Studies with ^{15}O-Water. Three groups of paired datasets were created from one dynamic study. The datasets in the three groups contained 400,000, 800,000 and 1.5 million counts. Synthetic reslicing was performed and repeated 50 times for each group.

RESULTS

All results are presented as the mean of the absolute values of the deviations from the correct values ± 1 s.d. of the absolute values. The reason for choosing the mean of the absolute values is that the mean of the deviations over 50 trials invariably differs very little from zero and carries little information about the error size. As a rule, results are presented as tables rather than figures of estimated values versus true values. The number of figures needed to present all the data would be very large and they carry no significant extra information compared to the tables. Maximum values are not presented for deviations since in the more than one thousand realignments that have been done in preparation for this paper there has not been a single outlier and no maximum value has deviated from what would be expected from the mean and the s.d.

The Effects of Initial Resampling of the Reference Image Set. Figure 4 demonstrates the effect of initial resampling of the reference image set. The double peaks apparent in the correlation and the SAD curve in Figure 4A disappear in Figure 4B where resampling was performed.

The Effects of Prefiltering of the Image Sets. The results from the experiments with different FWHMs on the filter used for data prefiltering before reorientation are presented in Tables 1 and 2 and in Figure 5. Figure 5 demonstrates that filter width has little effect on the registration errors in the study with good statistics. For the study with poor statistics, there is marked dependence on filter width and there seems to be a distinct minimum for a filter with a 20 mm FWHM.

TABLE 2
Registration Errors Versus Filter FWHM for Low Count Studies

FWHM	Translation (mm)			Rotation (deg)		
	x	y	z	α	β	γ
0	0.17 ± 0.14	0.21 ± 0.15	0.32 ± 0.19	0.22 ± 0.17	0.22 ± 0.18	0.63 ± 0.22
10	0.10 ± 0.11	0.26 ± 0.13	0.20 ± 0.11	0.10 ± 0.15	0.26 ± 0.34	0.36 ± 0.11
20	0.17 ± 0.12	0.19 ± 0.11	0.12 ± 0.10	0.08 ± 0.08	0.12 ± 0.09	0.18 ± 0.10
30	0.31 ± 0.18	0.19 ± 0.13	0.13 ± 0.12	0.09 ± 0.10	0.26 ± 0.19	0.28 ± 0.14
40	0.56 ± 0.25	0.28 ± 0.21	0.11 ± 0.17	0.06 ± 0.14	0.55 ± 0.24	0.56 ± 0.24

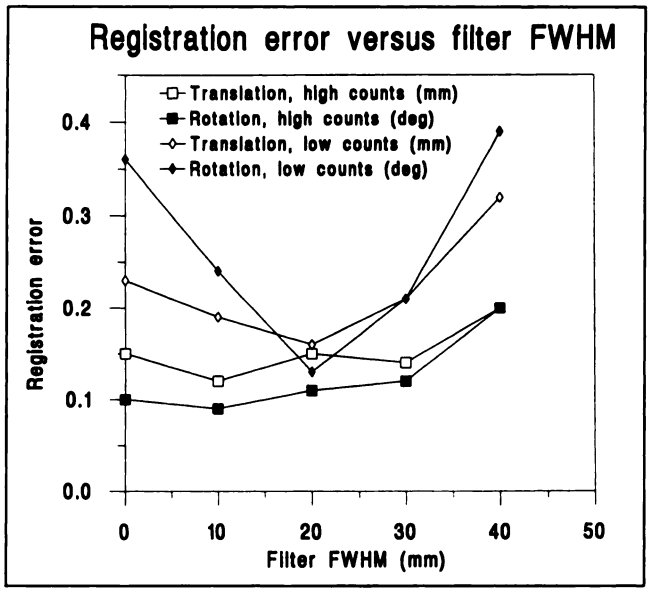

FIGURE 5. Registration error versus filter FWHM for high count (5.6 million counts) and low count (800,000 counts) studies. Translations and rotations in all directions are lumped and curves represent a mean of all directions.

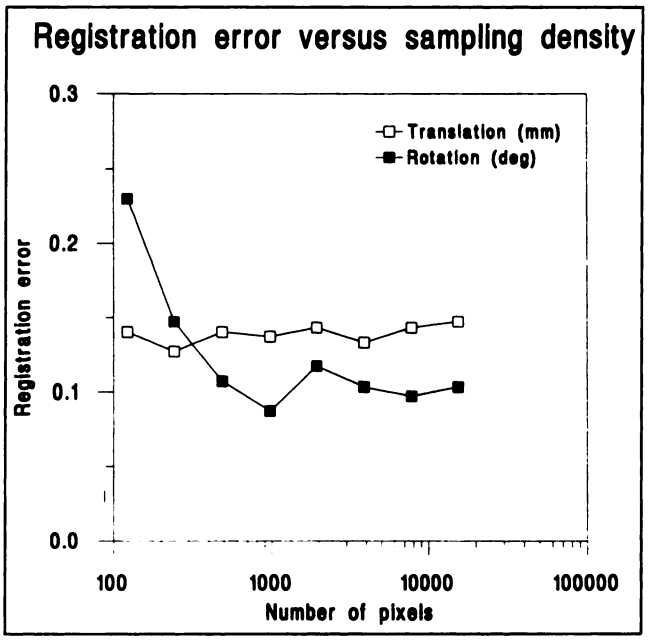

FIGURE 6. Registration error versus number of pixels included in the evaluation of the correlation coefficient. Translations and rotations in all directions are lumped and curves represent a mean of all directions.

Sampling Demands. Table 3 and Figure 6 show how registration errors change as the pixels within the mask undergo further subsampling. Figure 6 demonstrates how registration accuracy still remains unchanged when no more than 1 of 32 pixels (constituting a total of 486 pixels) within the mask are used.

Noise Sensitivity. Registration accuracy remains unchanged in the interval from 5.6 million counts per image set to 800,000 counts per image set, as demonstrated by Table 4 and Figure 7. When total counts decrease to 400,000 (equivalent to 5 sec on a CBF study), registration errors start to increase slightly.

Registration Accuracy for CBF Studies with and without Local Perturbations. Results from the synthetic reslicing of paired CBF studies are presented in Table 5. The simulated activation had little effect on registration errors while the simulated media infarction had a larger impact. Results from multiple realignment are presented in Table 6 and show the actual accuracy spread. Figure 8 shows the relation between true and estimated translation and rotation for the nonperturbed case and for the media infarction case.

Registration Accuracy for CBF Studies to Oxygen Studies. The results from realignment of oxygen studies to CBF studies are presented in Table 7.

Registration Accuracy for Brain Examinations with Markedly Different Tracers. Table 8 shows the results from the registration of image sets obtained with different tracers to each other. The errors are markedly larger compared to images obtained with the same tracer.

Registration Accuracy for Myocardial Studies with or without Perfusion Defects. The results from the registration of myocardial studies are shown in Table 9. Note that the accuracy in the acetate studies is comparable to that in the CBF studies despite a pixel size which is in the twice as

TABLE 3
Registration Errors Versus Sampling Density

Sampling	Translation (mm)			Rotation (deg)		
	x	y	z	α	β	γ
128	0.13 ± 0.12	0.15 ± 0.17	0.14 ± 0.18	0.23 ± 0.23	0.30 ± 0.35	0.17 ± 0.16
64	0.12 ± 0.09	0.15 ± 0.10	0.11 ± 0.10	0.16 ± 0.09	0.17 ± 0.11	0.11 ± 0.08
32	0.14 ± 0.10	0.17 ± 0.11	0.11 ± 0.10	0.13 ± 0.10	0.13 ± 0.11	0.06 ± 0.06
16	0.15 ± 0.11	0.17 ± 0.12	0.09 ± 0.07	0.09 ± 0.08	0.10 ± 0.09	0.07 ± 0.07
8	0.15 ± 0.11	0.17 ± 0.12	0.11 ± 0.08	0.13 ± 0.10	0.14 ± 0.09	0.08 ± 0.07
4	0.13 ± 0.10	0.17 ± 0.10	0.10 ± 0.08	0.13 ± 0.08	0.11 ± 0.09	0.07 ± 0.06
2	0.16 ± 0.11	0.16 ± 0.09	0.11 ± 0.09	0.12 ± 0.09	0.10 ± 0.09	0.07 ± 0.07
1	0.15 ± 0.11	0.18 ± 0.12	0.11 ± 0.08	0.12 ± 0.08	0.11 ± 0.10	0.08 ± 0.06

TABLE 4
Registration Errors Versus Total Number of Count in Studies

Counts	Translation (mm)			Rotation (deg)		
	x	y	z	α	β	γ
387,000	0.40 ± 0.21	0.16 ± 0.16	0.11 ± 0.16	0.39 ± 0.21	0.42 ± 0.23	0.79 ± 0.25
807,000	0.19 ± 0.12	0.18 ± 0.12	0.12 ± 0.12	0.11 ± 0.19	0.13 ± 0.21	0.21 ± 0.10
1,340,000	0.18 ± 0.11	0.17 ± 0.17	0.25 ± 0.15	0.19 ± 0.24	0.15 ± 0.12	0.15 ± 0.09
2,412,000	0.21 ± 0.14	0.14 ± 0.14	0.22 ± 0.13	0.21 ± 0.20	0.35 ± 0.21	0.15 ± 0.09
3,913,000	0.15 ± 0.11	0.13 ± 0.11	0.11 ± 0.08	0.11 ± 0.08	0.33 ± 0.15	0.15 ± 0.10
5,597,000	0.10 ± 0.09	0.21 ± 0.14	0.11 ± 0.16	0.16 ± 0.20	0.13 ± 0.10	0.08 ± 0.07

large in these studies. The introduction of a perfusion defect increased the registration error, but there was no error larger than 1.5 mm or 1.5 degrees in any direction in the 50 reslicings we performed.

DISCUSSION

The explanation for the double peaks in Figure 4A is that when comparing the two datasets the correlation is lowered by the presence of noise. When one of the datasets is moved in the z-direction, the resulting interpolation has a considerable noise reducing effect. Thus, when the sample dataset is moved in the z-direction, the correlation diminishes due to increased misregistration, but also increases due to smoothing from the interpolation. Initially, smoothing affects the correlation more than misregistration and there will be two peaks symmetrically placed around the true minimum. This interpretation is supported by the finding that this effect increases (the peaks get higher and further apart) with a decreased number of counts and decreases with increased filter FWHM. A remedy for this, as seen in Figure 4B, is to resample the reference dataset into a volume with a different axial sampling, where the new sampling density is not an even multiple or fraction of the initial sampling. Thus, every reslicing of the sample dataset will produce the same type of noise reduction regardless of whether or not it includes any z-translations. Another solution would be to resample both datasets initially so that the voxel size in the z-direction is small compared to the FWHM (*2*), this would be more time-consuming. The problem may not be detected by experimental setups where there are actual out-of-plane translations or rotations in the datasets used to verify realignment accuracy. Neglect of the problem will result in systematic z-direction misalignments of 0.5–1.0 mm on correctly aligned data.

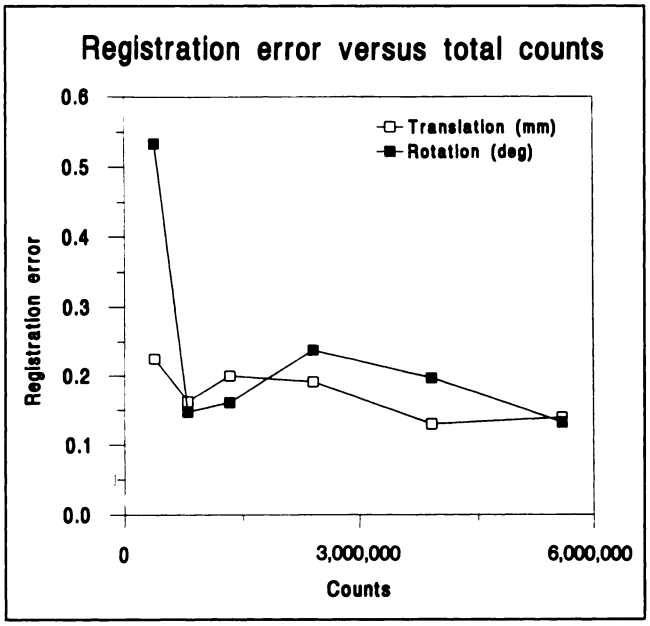

FIGURE 7. Registration error versus total number of counts per 15 slices. Translations and rotations in all directions are lumped and curves represent a mean of all directions.

Different filter widths seem to have little influence on realignment accuracy on studies with good statistics (total counts 5.6 million) as evident from Figure 5. This is in agreement with findings of Eberl et al. (*5*), but in contrast to Woods et al. (*2*) who report a decrease in registration error with increased filter FWHM. On the low count dataset (total counts 800,000), on the other hand, there is a marked dependence on filter width with a distinct minimum of 20 mm FWHM; consequently, that is the filter width of choice for the rest of the study.

Figure 6 shows that it is feasible to subsample the volume so that less than 0.5% of the data (486 voxels) is actually used in the fitting procedure with no loss of precision. The reason for this, in part, is that with the heavy filtering applied to the data, every voxel really contains information from a large number of adjacent voxels. Thus, if every voxel were to be used, the same data would be used over and over again, adding no real information. This is in accordance with calculations showing that for data filtered to obtain a FWHM of 20 mm there are no more than roughly 500 independent information elements (*15*) in a 15-slice volume.

The data discussed above have led to the following protocol for the practical use of the method. The datasets are filtered with a Gaussian filter with a FWHM of 20 mm. The reference dataset is resampled in the axial direction from the initial 15 slices of 6.5 mm thickness to 19 slices with a thickness of 5.0 mm. The reference volume is differentiated and a mask is generated from the five percent of the voxels with the highest derivative values. The dataset is further subsampled so that every 64th voxel is used for the first full

TABLE 5
Registration Errors for CBF Studies

Translation (mm)			Rotation (deg)		
x	y	z	α	β	γ
0.10 ± 0.09	0.21 ± 0.14	0.11 ± 0.16	0.16 ± 0.20	0.13 ± 0.10	0.08 ± 0.07
0.16 ± 0.19*	0.18 ± 0.13	0.10 ± 0.11	0.14 ± 0.19	0.16 ± 0.20	0.09 ± 0.08
0.97 ± 0.35†	0.11 ± 0.20	0.90 ± 0.32	0.66 ± 0.29	0.60 ± 0.38	0.72 ± 0.31

*One of the sets contained a simulated activation effect.
†One of the sets contained a simulated middle cerebral infarction.

Powell iteration, every 32nd voxel for the second and every 16th voxel for the subsequent iterations.

The noise sensitivity of the method is an important parameter that has been insufficiently investigated in previous studies when registering PET images. Hoh et al. (*3*) report their results on data with 6.7 million counts per dataset. Eberl et al. (*5*) compared alignment accuracy for datasets acquired for 1 min and 15 min and found no difference. Judging from the amount of activity they injected into their phantom, however, the total number of counts should still be several million for the 1-min dataset and cannot really be considered a low count study. Woods et al. (*2*) reported their results on data containing roughly the same amount of counts as a normal CBF study. The latter group indicated that the registration errors increased when the number of counts were reduced by half. Results from our study, on the other hand, indicate that registration errors essentially remain the same down to 800,000 counts. This is comparable to what is collected for 5–10 sec during a CBF study.

There is no evidence from our data that the method should function worse in one direction than another. No general difference in accuracy is found between translations or rotations in the different directions. This is in contrast to results reported by Hoh et al. (*3*) who obtained better precision for in-plane translations and rotations than out-of-plane translations and rotations. Neither the poorer intrinsic sampling density in the axial direction should be of any consequence, considering the results from the experiment with different sampling frequencies which show that it is possible to subsample the datasets in the x and y directions once every 8 mm with no loss of accuracy, nor should the marginally poorer resolution in the axial direction, given the results from the experiments with different filter FWHM. Actually, for the filter width (FWHM 20 mm) used in the subsequent experiments in this article, the resolution was more than twice as good in the axial as in the transaxial direction.

Figures 8A and B show estimated translations and rotations versus the true induced values for a CBF study with 5.6 million counts. The good agreement is typical of results obtained when realigning brain or cardiac studies using the same tracer against each other. Figures 8C and D show the results when realigning a CBF dataset with a media infarction versus a nonperturbed CBF dataset. The trend apparent in these graphs with increased absolute errors and relatively unchanged spread of the errors is similar to what was found when realigning cardiac studies with local defects and when realigning studies obtained with different tracers against each other.

There is a marked degradation in performance of the method when applied to data with large focal defects. Yet, registration is well within the size of one voxel. The results compare well to the findings of Hoh et al. (*3*) who compared the results for SAD and SSC on image sets with focal defects and found them to be essentially equal despite earlier claims that SAD is inferior for dissimilar images (*8*).

As an implementation detail, the software developed to perform the registration includes the possibility for the user to interact and explicitly exclude certain parts of the image sets from the analysis. This is done by drawing an arbitrary number of regions of interests (ROIs) outlining the areas of the dataset to be excluded from further analysis. When the final mask is created, there is a NAND function between the mask created from the derivatives and from the ROIs. When this option is used for brain and cardiac datasets containing focal defects, the same registration accuracy as that in datasets without focal defects was achieved. This type of amendment would be feasible for use in any of the previously published methods (*1–5*) and is not specific to the method presented here.

The ability to realign PET examinations performed with

TABLE 6
Registration Errors from Multiple Realignment for CBF Studies

Translation (mm)			Rotation (deg)		
x	y	z	α	β	γ
0.11 ± 0.11	0.18 ± 0.15	0.13 ± 0.08	0.18 ± 0.15	0.15 ± 0.13	0.18 ± 0.15

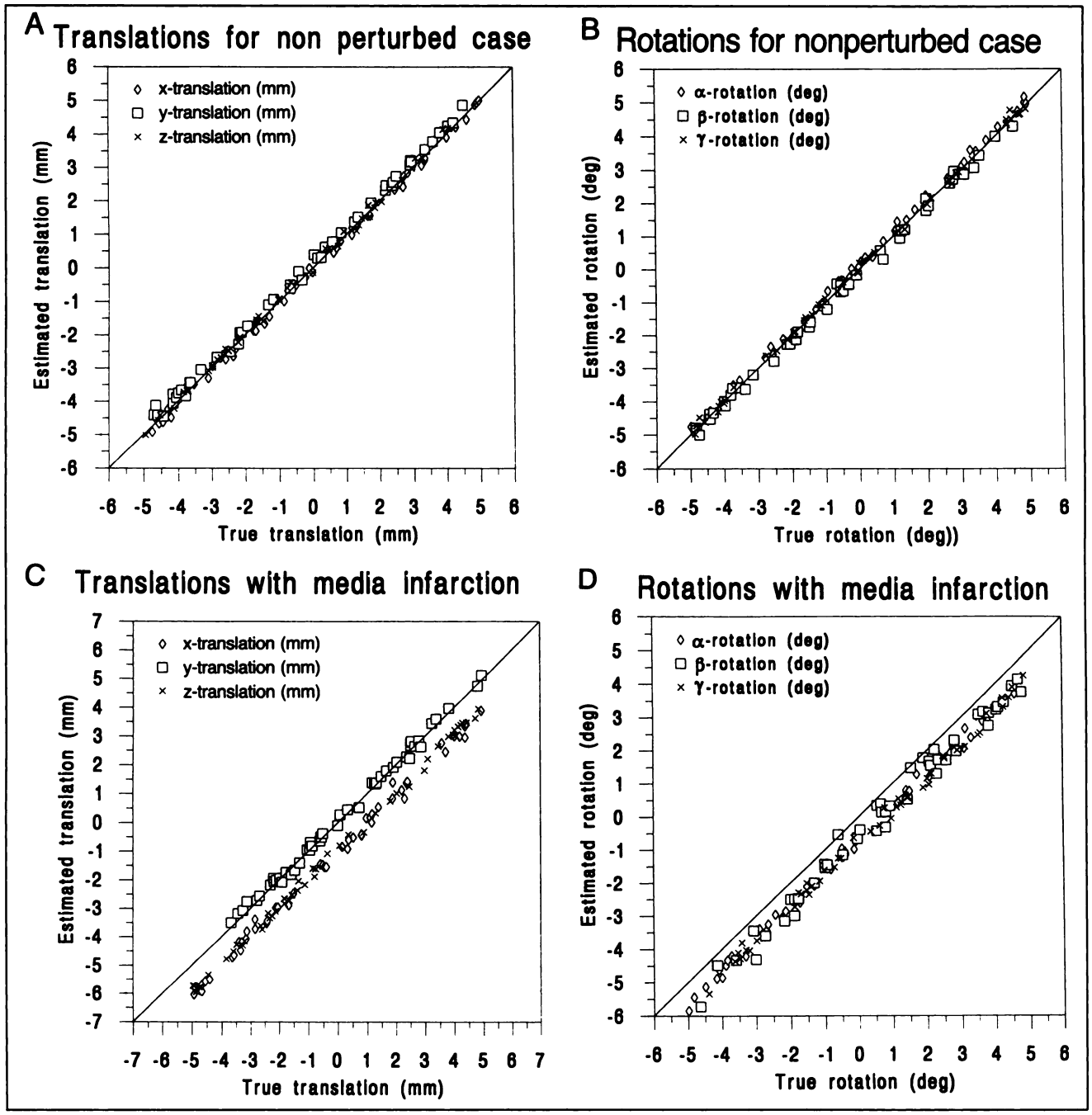

FIGURE 8. A and B show estimated translations and rotations versus the true induced ones for CBF studies containing 5.6 million counts. C and D show the results when one of the studies contain a large simulated middle cerebral infarction.

different tracers may well prove useful. Movements between the oxygen and the CBF studies have been a problem at our facility when measuring the oxygen extraction ratio (OER) and metabolism. The need to administer oxygen through a mouthpiece makes it difficult to achieve good fixation, and the subjects in this type of examination often

TABLE 7
Registration Errors for Oxygen-15-Water Registration

Translation (mm)			Rotation (deg)		
x	y	z	α	β	γ
0.24 ± 0.14	0.23 ± 0.12	0.58 ± 0.21	0.43 ± 0.16	0.26 ± 0.20	0.14 ± 0.12

TABLE 8
Registration Errors for Different Tracers

Tracer	Translation (mm)			Rotation (deg)		
	x	y	z	α	β	γ
NMSP-DOPA	0.12 ± 0.10	0.25 ± 0.25	1.53 ± 0.58	1.00 ± 1.18	0.50 ± 0.55	0.58 ± 0.22
NMSP-H_2O	0.09 ± 0.11	1.25 ± 0.22	0.86 ± 0.25	0.22 ± 0.18	0.51 ± 0.28	1.15 ± 0.21
H_2O-DOPA	0.13 ± 0.11	1.14 ± 0.24	0.61 ± 0.36	1.27 ± 0.54	0.51 ± 0.46	0.47 ± 0.21

have difficulty in cooperating [for example, patients with acute ischemic stroke (*10*)]. The effects of movement between the oxygen and the CBF study may be subtle and resemble a localized increase in OER. Therefore, the technique described in this paper is valuable in improving the quality of clinical and scientific measurements of oxygen extraction and metabolism.

The ability to register examinations with widely varying regional uptake patterns in ^{15}O-water, ^{11}C-DOPA and NMSP facilitate the evaluation of multitracer protocols and opens the way for multivariate pixel-by-pixel analysis of such datasets.

The results for realignment of scans obtained with different tracers are significantly worse than when the same tracer is used. This may be partially explained by actual movement between the scan sessions. The individual examined with labeled DOPA, NMSP and water was lightly sedated but not muscle relaxed, and since the DOPA and NMSP examinations were 2 hr apart, it is possible that some movement actually occurred. When validating registration in this manner, assuming no movement between two experiments, the results are always worst case.

The technique of extracting parts of the object with a high signal-to-noise ratio with respect to orientational information by the use of derivative information can be viewed as an intermediate between surface fit (*9*) and image similarity techniques. The result is a method that will find and use the information that exists in sharp edges, between, for example, the brain or myocardium and surrounding tissues, in the cases where it exists, but will degenerate to an ordinary image similarity technique, performing no worse or better when there are no sharp edges. Compared to ordinary surface fit techniques, this method has the advantages of automatically using edges inside the outer contour and is easy to implement. When contours from internal structures have been used in the surface fit technique, they have been entered manually (*16*). The surface fit technique, on the other hand, has the advantage of being designed for intermodality registration (*9*) and has been successfully applied to whole-body intermodality registration (*16*).

Further research in this area should concentrate on the effects of misalignment between transmission and emission scans. In a typical setting, one transmission scan is obtained followed by multiple emission scans. These emission scans are then realigned with the first emission scan. This means that if there are misalignments in the emission datasets, they will all have been reconstructed with different positioning of the transmission relative to the emission dataset. It is clearly not meaningful to discuss registration errors in tenths of millimeters from initial misalignments of 10 mm or more if enough attention is not paid to obtaining correct attenuation correction. Experiments performed at our facility indicate that translations of 5 mm between the transmission and emission brain study may cause local quantification errors of up to 10%. Work is currently being done to develop a method to correct these effects.

CONCLUSIONS

The use of edge information from image volume provides a rapid, accurate and noise-insensitive method to realign PET datasets. This method has been applied to data with no more than 800,000 counts per 15 slices and normally requires about 1 min of CPU time on the computer.

TABLE 9
Registration Errors in Cardiac Studies versus Tracer and Total Number of Counts in the Study

Tracer	Counts	Translation (mm)			Rotation (deg)		
		x	y	z	α	β	γ
Acetate	4,357,000	0.14 ± 0.11	0.10 ± 0.07	0.21 ± 0.16	0.26 ± 0.19	0.12 ± 0.10	0.14 ± 0.09
Acetate	8,761,000	0.16 ± 0.12	0.10 ± 0.07	0.16 ± 0.11	0.26 ± 0.18	0.13 ± 0.11	0.10 ± 0.09
Acetate*	4,357,000	0.95 ± 0.19	0.53 ± 0.10	0.34 ± 0.25	0.21 ± 0.17	0.95 ± 0.23	0.40 ± 0.19
H_2O	424,000	0.25 ± 0.10	0.27 ± 0.10	0.54 ± 0.15	0.56 ± 0.13	0.20 ± 0.13	0.75 ± 0.14
H_2O	836,000	0.09 ± 0.08	0.08 ± 0.09	0.22 ± 0.13	0.53 ± 0.12	0.14 ± 0.10	0.37 ± 0.11
H_2O	1,540,000	0.08 ± 0.08	0.09 ± 0.06	0.24 ± 0.13	0.41 ± 0.11	0.13 ± 0.10	0.41 ± 0.12

*One of the datasets contained a large simulated perfusion defect in the left chamber.

This enables its routine use for realignment of individual frames in dynamic sequences. The method has also been shown to be capable of realigning datasets obtained with different tracers with markedly dissimilar regional uptake, thus facilitating the use of pixel-by-pixel techniques in multitracer protocols.

APPENDIX

To indicate the gains in execution time realized by using the edge information, estimates of the execution time for the various parts of the process are given. The time estimates are based on measurements of CPU time from 50 different runs with different initial misalignments. Time estimates for I/O (reading and writing images to and from disc) are not given.

1. Read Images. Read the two image sets and denote one of them reference volume and the other sample volume. A representative slice from a CBF study containing 2.5 million counts is shown in Figure 9A.

2. Prefilter Images. A normalized Gaussian profile with a given FWHM is created. The profile is truncated when the amplitude goes below 30% of the peak value. A profile with a 10 mm FWHM to be applied to an image with a 2 mm pixel size would then look as follows; {0.17, 0.37, 0.64, 0.90, 1.00, 0.90, 0.64, 0.37, 0.17}. Since the gaussian is separable in the x- and y-directions, the filtering is performed as two one-dimensional convolutions with the kernel seen above. Thus the filtering will be a 2N rather than an N^2 process (where N is the size of the kernel), saving some computation time. The total time to filter two image sets of 15 slices each is approximately 17 sec. The same slice as in Figure 9A is shown in Figure 9B after filtering with a 20 mm filter.

3. Reslice the Reference Volume. The reference volume is resliced with a changed axial sampling. The original image volume consists of 15 slices with a slice distance of 6.5 mm, and is resliced to 19 slices with 5 mm slice distance. Thus, after reslicing only the middle slice (slice 8 in the original and 10 in the resliced volume) will coincide exactly. The effect of this is that when the sample volume is resliced the same noise reduction due to averaging over original slices is achieved regardless of whether any out-of-plane translations or rotations are applied or not. As is explained in the main text, this eliminates the problem of artifactual z translations. The actual reslicing algorithm is described in point seven. The total time needed to resample the reference volume to 19 slices is approximately 10 sec.

4. Differentiate the Reference Volume. A one-dimensional differentiating filter kernel ({−1.0/pixel size, 0.0, 1.0/pixel size}) is created for each of the x, y and z directions. These filters are applied to the original image volume yielding three new volumes (Fig. 9C, D, E) depicting the magnitude of the derivative in each of the directions. The volumes represent vector fields and are added as such (taking the square root of the sum of the squared partial derivatives) to yield values of the magnitude of the total derivative (Fig. 9F). It is evident from that figure that the derivatives and the information content with respect to orientation is indeed not homogeneous and that the cortical surface and the borders between gray and white matter have the highest information content. The kernel described above will yield values that are to some extent contaminated with higher order derivatives. Tests with more elaborate kernels (and thus larger and more time consuming) to minimize these contributions have shown that this is of no practical consequence. The time needed for differentiation and summation of derivatives is approximately five seconds.

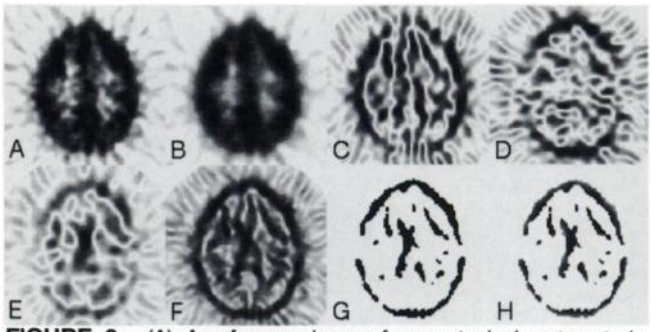

FIGURE 9. (A) A reference image from a typical water study containing 2.5 million counts per 15 slices. (B) Reference image after filtration to 20 mm FWHM. (C) Resulting image after differentiation in the x-direction. (D) Resulting image after differentiation in the y-direction. (E) Resulting image after differentiation in the z-direction. (F) Image of total derivative. (G) Mask created by tresholding derivative image. (H) Sample image resliced within mask.

5. Threshold the Derivative Volume. The derivative volume is scanned for its maximum value and a histogram is created with 1000 bins between zero and the maximum value, where the value of each bin represents the number of pixels that have a value within the range represented by that bin. The histogram is traversed from the maximum value and down, summing the values, until the sum exceeds five percent of the total number of pixels in the volume. The exact threshold value is obtained by interpolation between adjacent bins. A slice of the resulting mask volume is shown in Figure 9G. To evaluate the threshold and create a mask takes approximately one to two seconds.

6. Choose Initial Transformation Parameters. Zero values are used as the initial guess for both translations and rotations.

7. Reslice the Sample Set. The routine to reslice a volume proceeds in the following way:

a. Create an empty volume of the same dimensions as the reference volume.
b. Create a 4 × 4 transformation matrix from the values for translation and rotation.
c. Set index in empty volume to {0,0,0} and call the index {xi,yi,zi}.
d. If the mask is zero in this position go to j, else proceed.
e. Obtain coordinates from index in empty volume using information on pixel size in reference volume in all three dimensions and set fourth value to 1, yielding coordinates {x,y,z,1.0}.
f. Multiply the transformation matrix with coordinates, yielding {x′,y′,z′,1.0}, that is the coordinates in the sample volume.
g. Obtain index from coordinates in sample volume using information on pixel size in sample volume in all three directions, yielding {xi′,yi′,zi′}. By using truncation when proceeding from coordinates to index, the index of the pixel with the lowest value in all three directions of the eight pixels surrounding the point {x′,y′,z′} will be obtained.
h. Perform a trilinear interpolation between the eight surrounding points using the information on difference in position between {x′,y′,z′} and {xi′,yi′,zi′} and write the resulting value into the empty volume in position {xi,yi,zi}.
i. Step in empty volume by increasing {xi,yi,zi} in one direction.
j. If entire volume has been traversed stop, else go to d.

By incorporating the possibly different pixel sizes in the reference and the sample volumes, the same routine may be used to simply resample a volume to a new pixel size by setting translations and rotations to zero. This is how it is used to accomplish the resampling in the z direction described in point three above. An example of a sample volume that has been resliced within the corresponding mask is shown in Figure 9H. The time needed to reslice a volume clearly depends on the level of further subsampling that is used. If a subsampling level of every eighth pixel is used (the level that is recommended in the main text for the final iteration) a reslicing finishes in approximately 0.2 sec if a mask is used and in 2.9 sec if not. The full time saving of a factor of 20 is not realized since the check for one or zero in the mask also requires some time.

8. Evaluate the Correlation. The correlation coefficient is evaluated according to the equation:

$$CC = \frac{N \sum x_i y_i - \sum x_i y_i}{\sqrt{N \sum x_i^2 - \left(\sum x_i\right)^2} * \sqrt{N \sum y_i^2 - \left(\sum y_i\right)^2}},$$

where x_i and y_i denotes pixel values in the reference and the sample volumes respectively, where N is the number of non-zero values in the mask volume and where all summations are performed over pixels with non-zero values in the mask.

9. Check for Convergence. The Powell method proceeds by carrying out subsequent line minimizations in different directions in the parameter space, that is for this application six-dimensional. One iteration in the Powell method thus consists of six line minimizations, each one consisting of a number of function evaluations, where a function evaluation consists of steps seven and eight above. Since they are at the core of the iterative process the number of times that these steps are executed typically becomes very large (200 to 400 times) which is why any attempt to increase execution speed should be directed towards increasing their efficiency. Convergence for each line minimization is said to have been achieved when the points bracketing the minimum are less than 0.01 mm or degrees apart, and convergence for the entire minimization is signaled by the failure to improve on the correlation coefficient by more than 0.01 percent in one full iteration. It has been found that due to the intrinsics of the implementation of the Powell method that is used (it uses golden section to first bracket and then to narrow down on the minimum) in the line minimizations the same points in the parameter space are frequently visited more than once. This implies that the same computations will be performed many times. This has been addressed by a cache where translations, rotations and the resulting correlation coefficient is stored, and every time a new point in the parameter space is to be evaluated, the cache is first searched to see if the information is already available. The time needed for the full iterative procedure varies somewhat from case to case depending on magnitude of initial misalignment. For the sub-sampling strategy outlined in the main text, typical times range from 15 to 25 sec using a mask and from 350 to 500 sec without. Note that here there is a factor of 20 in time gain despite a smaller gain for the reslicing. Since the CC is considerably more peaked when using a mask the convergence criteria affects the two cases somewhat differently, explaining why there are a larger number of iterations when not using a mask, in addition to each iteration requiring more time.

10. Reslice the Sample Volume. The sample volume is resliced using the values for translation and rotation that were found to give the highest value for the CC. The reslicing uses approximately eight seconds of CPU time.

11. Write Resliced Volume. The resliced sample volume is written to a file.

The total execution time for the full procedure is approximately one minute when using the mask, and around six to nine minutes if without. This time estimate is exclusive of disc I/O time. It should be pointed out that as the speed of the iterative procedure is increased, the relative importance of the other parts, and also of the I/O part, increases. The source code of the software described above can be obtained, free of charge, by written request to the author.

ACKNOWLEDGMENTS

The author thanks Johan Bosaeus, Anders Lilja, Sven Valind and Bengt Långström for their valuable discussions of the work and constructive criticism of the manuscript.

REFERENCES

1. Mintun MA, Lee KS. Mathematical realignment of paired PET images to enable pixel-by-pixel subtraction [Abstract]. *J Nucl Med* 1990;31(suppl): 816.
2. Woods RP, Cherry SR, Mazziotta JC. Rapid automated algorithm for aligning and reslicing PET images. *J Comput Assist Tomogr* 1992;16:620–633.
3. Hoh CK, Dahlbom M, Harris G, et al. Automated iterative three-dimensional registration of positron emission tomography images. *J Nucl Med* 1993;34:2009–2018.
4. Bacharach SL, Douglas MA, Carson RE, et al. Three-dimensional registration of cardiac positron emission tomography attenuation scans. *J Nucl Med* 1993;34:311–321.
5. Eberl S, Kanno I, Fulton RR, et al. Automatic three-dimensional spatial alignment for correcting interstudy 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 brain PET*. Elsevier Science Publishers B.V., Amsterdam; 1993:419–426.
6. Bergström M, Boëthius J, Eriksson L, Greitz T, Ribbe T, Widén L. Head fixation device for reproducible position alignment in transmission CT and positron emission tomography. *J Comput Assist Tomogr* 1981;5: 136–141.
7. Evans AC, Marrett S, Torrescorzo J, Ku S, Collins L. MRI-PET correlation in three dimensions using a volume-of-interest (VOI) atlas. *J Cereb Blood Flow Metab* 1991;11(suppl):A69–A78.
8. Venot A, Lebruchec JF, Roucayrol JC. A new class of similarity measures for robust image registration. *Comput Vis Graph Image Process* 1984;28: 176–184.
9. Pelizzari CA, George TYC, 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;13:20–26.
10. Heiss WD, Huber M, Fink GR et al. Progressive derangement of periinfarct viable tissue in ischemic stroke. *J Cereb Blood Flow Metab* 1992;12:193–203.
11. Powell MJD. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *J Comput Assist Tomogr* 1964;7:155–163.
12. Press WH, Flannery BP, Teukolsky SA, Wetterling WT. *Numerical recipes in C. The art of scientific computing*. New York: Cambridge University Press; 1988:291–317.
13. Holte S, Eriksson L, Dahlbom M. A preliminary evaluation of the Scanditronix PC2048-15B brain scanner. *Eur J Nucl Med* 1989;15:719–721.
14. Kops R, Herzog H, Schmid A, Holte S, Feinendegen L. Performance characteristics of an eight ring whole-body PET scanner. *J Comput Assist Tomogr* 1990;14:437–445.
15. Worsley KJ, Evans AC, Marret S, Neelin P. A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 1992;12:900–918.
16. Scott AM, Macapinlac H, Zhang JJ, et al. Clinical applications of fusion imaging in oncology. *Nucl Med Biol* 1994;5:775–784.