NEUROIMAGE **8,** 30–43 (1998)
ARTICLE NO. NI980335

Registration of MR/MR and MR/SPECT Brain Images by Fast Stochastic Optimization of Robust Voxel Similarity Measures

Christophoros Nikou,*,† Fabrice Heitz,† Jean-Paul Armspach,* Izzie-Jacques Namer,* and Daniel Grucker*

*Faculté de Médecine, Institut de Physique Biologique, 4 rue Kirschleger, 67085 Strasbourg, France; and †Laboratoire des Sciences de l'Image, de l'Informatique, et de la Télédétection, 4, Boulevard Sébastien Brant, 67400 Illkirch, France

Received August 6, 1997

This paper describes a robust, fully automated algorithm to register intrasubject 3D single and multimodal images of the human brain. The proposed technique accounts for the major limitations of the existing voxel similarity-based methods: sensitivity of the registration to local minima of the similarity function and inability to cope with gross dissimilarities in the two images to be registered. Local minima are avoided by the implementation of a stochastic iterative optimization technique (fast simulated annealing). In addition, robust estimation is applied to reject outliers in case the images show significant differences (due to lesion evolution, incomplete acquisition, non-Gaussian noise, etc.). In order to evaluate the performance of this technique, 2D and 3D MR and SPECT human brain images were artificially rotated, translated, and corrupted by noise. A test object was acquired under different angles and positions for evaluating the accuracy of the registration. The approach has also been validated on real multiple sclerosis MR images of the same patient taken at different times. Furthermore, robust MR/SPECT image registration has permitted the representation of functional features for patients with partially complex seizures. The fast simulated annealing algorithm combined with robust estimation yields registration errors that are less than 1° in rotation and less than 1 voxel in translation (image dimensions of 128^3). It compares favorably with other standard voxel similarity-based approaches. © 1998 Academic Press

Key Words: **magnetic resonance imaging (MRI); single photon emission computed tomography (SPECT); multimodal registration; stochastic optimization; robust estimation.**

INTRODUCTION

The goal of medical image matching is to geometrically align two or more image volumes or surfaces so that voxels representing the same anatomical structure may be superimposed. During the past two decades, the progress in neuroimaging has revolutionized clinical research in neurology and neurosurgery. Modern techniques provide structural (MR imaging), functional (SPECT, PET, functional MR imaging), or metabolic (PET, MR spectroscopy) information considered vital nowadays, not only in order to understand the physiopathology of many diseases, but also for diagnosis and evaluation of treatment efficacy. Repeatedly acquired MR and SPECT images from the same patient motivate the detection of changes. From this perspective, good-quality registrations are required.

A large variety of image registration methods have been proposed for medical applications. A general review has been made by Brown (1992) and a classification presented by Van den Elsen *et al.* (1993). A first category of registration methods implies human interaction (Pietrzyk *et al.,* 1994; Turkington *et al.,* 1995). In this paper we are concerned with fully automated (or data-driven) methods, relying on more sophisticated image models and algorithms. "Similarity measure-based approaches" rely on the minimization of cost functions that express the pixel or voxel similarity of the images to be aligned. Similarity measures have been proposed for both single and multimodal medical image registration. In the case of single modal image registration, the cost function is generally related to a Gaussian sensor model (Christensen *et al.,* 1996) or, equivalently, to least-squares estimation (Woods *et al.,* 1992; Hajnal *et al.,* 1995; Alpert *et al.,* 1996; Van den Elsen *et al.,* 1995; Pelizzari *et al.,* 1989; Friston *et al.,* 1995). Other similarity measures, based on standard image statistics (mean, variance (Woods *et al.,* 1993) or entropy measures (Wells *et al.,* 1996)), have been devised in the case of multimodal medical images. Other image registration approaches include principal axes registration (Alpert *et al.,* 1990; Arata *et al.,* 1995; Dhawan *et al.,* 1995) or registration by deformable templates (Bajcsy and Kovačič, 1989; Siomka *et al.,* 1995; Davatzikos, 1996).

Similarity measure-based approaches have been shown to be very efficient for coping with standard

1053-8119/98 $25.00
Copyright © 1998 by Academic Press
All rights of reproduction in any form reserved.

registration problems but suffer from several shortcomings that may be detrimental in specific registration situations. At first, cost functions related to standard similarity measures are generally highly nonlinear, yielding many local minima in the cost function. As a consequence, registration algorithms must be initialized close to the optimal (or desired) solution in order to obtain satisfactory solutions.

A second important limitation of existing algorithms is related to the underlying sensor models. For the single model case, the Gaussian sensor model assumes that the two images to be aligned differ only by additive Gaussian noise (Christensen *et al.*, 1996). This does not account for several standard situations in image registration such as incomplete acquisition or lesion evolution in an image sequence. In the multimodal case the most frequently used method is based on the image uniformity cost function introduced for MR/PET image registration (Woods *et al.*, 1993). This similarity measure assumes that a uniform region in the MR image corresponds to a uniform region in the PET image. This is generally only a crude approximation, since multimodal images are precisely used for the complementary information they provide to the physician. In addition, due to Compton scattering in SPECT images, anatomical structures do not occupy the same volume in MRI and SPECT image sequences. Multimodal images may thus differ significantly, and these differences are not uniform. The standard similarity-based approaches (Woods *et al.*, 1992, 1993; Hajnal *et al.*, 1995; Alpert *et al.*, 1996; Van den Elsen *et al.*, 1995; Pelizzari *et al.*, 1989; Friston *et al.*, 1995) do not model these information differences (called innovation in the information theoretic literature) and, as a consequence, are not robust with respect to them. Standard methods may thus result in inaccurate registrations or even misregistrations.

In this paper we propose a similarity measure-based approach that addresses the above-mentioned shortcomings. Robust statistics are introduced in both single modal and multimodal image registration similarity measures in order to account for significant image differences (these differences are called "outliers" in the robust statistics literature). The sensitivity of the registration algorithm to local minima in the similarity measure is also reduced by using a multigrid, fast stochastic optimization algorithm. As a consequence, the proposed method needs no initialization near the desired solution and is able to register reliably 3D single (MR/MR) and multimodal (MR/SPECT) brain images exhibiting significant differences (due to lesion evolution, noise, incomplete images, etc.).

Cost functions for the registration of single modal images, which are to a certain extent robust to image changes, have been also described (Herbin *et al.*, 1989; Minoshima *et al.*, 1992). Herbin *et al.* made use of deterministic and stochastic sign change criteria to provide robust registrations in critical situations corresponding for instance to lesion evolution in 2D images, whereas Minoshima *et al.* improved the technique by proposing its 3D extension for registering PET images presenting lesion evolution. Another robust statistics-based approach has been proposed recently and independently (Alexander and Somorjai, 1996) for the registration of 2D single modality images. The approach described by Alexander and Somorjai (1996) is feature-based and relies on a least median of squares robust estimator (Rousseeuw, 1984). Contrary to the method described below, its application to 3D images is not straightforward, due to the high computational complexity of the least median of squares estimator.

The robust registration technique proposed here has been validated using a test object acquired under different angles and positions, with various noise statistics. The 3D registration of a patient's brain MR/SPECT images has also been compared to the manual registration provided by an expert physician. The robust registration method compares favorably with the other standard similarity measure-based approaches, such as the least-squares cost function (Hajnal *et al.*, 1995), the image uniformity cost function (Woods *et al.*, 1993), and the entropy-based similarity measure (Wells *et al.*, 1996). Subvoxel accuracies are obtained in all cases by the robust methods. Experimental results for real-world cases are also presented and commented on. Single modal robust registration is being used with success to follow the lesion evolution of multiple sclerosis patients in MRI sequences. The robust multimodality method has been applied to superimpose SPECT image data of patients with partial complex epilepsy onto the MRI of the same patients.

MATERIALS AND METHODS

Our approach combines a robust regression method with a fast, multigrid, stochastic optimization of the similarity measure. The optimization algorithm is applied on a multigrid sequence, using a top-down approach starting from the coarsest resolution level (Heitz *et al.*, 1994). The solution obtained at a given resolution level is interpolated and forwarded to the next, finer resolution. Begining from voxel [0, 0, 0] and scanning the volume from front to back, top to bottom, and left to right, the algorithm first carries out the calculations for every 81st voxel in the 3D images. After the algorithm has converged, the resulting registration parameters represent the initial estimate for the next level, where every 27th voxel is processed, then every 9th voxel, every 3rd, and finally every voxel in the MRI image volume. The above multigrid technique is applied without the images being smoothed.

Multigrid matching is usually motivated by the

significant computational gain obtained in 3D registrations. As noticed by several authors (Heitz *et al.*, 1994), multigrid approaches are also far less sensitive to local minima in the cost function than a standard single-resolution optimization scheme. This yields fast convergence toward optimum solutions. At a given resolution level, the registration algorithm may be described as follows:

1. Fast segmentation of the two volumes to be registered from their background.
2. Gray-level normalization of the two 3D images.
3. Preregistration of the volume centroids in order to provide an initial guess for the 3D translation parameters.
4. Registration of the volumes of interest by fast stochastic optimization of the robust similarity measure.
5. Fine-tuning of the solution using a deterministic optimization algorithm.
6. Interpolation of the 3D translation and rotation parameters to be forwarded to the next (finer) resolution level.

At first, the reference image and the image to be registered are segmented from their background by simple thresholding. We have applied a nonparametric unsupervised algorithm (Otsu, 1979) to eliminate the noisy background, followed by a simple region-growing algorithm.

In the multimodality case the brain of the MRI is extracted by thresholding (Otsu, 1979) followed by a region-growing and connected-component analysis.

The average gray level of the reference image $I_{\text{ref}}(\cdot)$ and of the image to be registered $I_{\text{reg}}(\cdot)$ are normalized according to

$$I_{\text{ref}}(x) := \frac{\bar{I}_{\text{reg}}}{\bar{I}_{\text{ref}}} I_{\text{ref}}(x), \qquad \forall x = (i, j, k), \tag{1}$$

where $\bar{I}_{\text{ref}}$ and $\bar{I}_{\text{reg}}$ denote the expected values of images $I_{\text{ref}}(\cdot)$ and $I_{\text{reg}}(\cdot)$, respectively. The number of gray levels of the two volumes is also normalized to the same number G of gray levels, typically $G = 256$, by rescaling image values between the threshold given for the brain structure (Otsu, 1979) and the maximum gray level of the image.

The centroids of the two normalized segmented 3D images are then registered in order to compensate for a large part of the 3D translation. The next step consists of estimating the parameters Θ^* of the rigid transformation T_Θ, minimizing a robust cost function $E(T_\Theta(x))$, that expresses the similarity between the single and the multimodal images (the cost functions used in our experiments are presented in the next section)

$$\Theta^* = \text{argmin}\,[E(T_\Theta(x))], \tag{2}$$

where

$$\Theta = (t_x, t_y, t_z, \hat{\theta}_x, \hat{\theta}_y, \hat{\theta}_z)^T$$

is a vector containing the 3D translation parameters (t_x, t_y, t_z) with respect to the x, y, and z axes and the Euler rotation angles $(\hat{\theta}_x, \hat{\theta}_y, \hat{\theta}_z)$. In our case, the scaling factors are calculated by the known voxel sizes as provided by the acquisition systems. However, their estimation can be easily integrated in the proposed algorithm.

Fast stochastic optimization and deterministic, gradient-descent algorithms are used to perform this optimization step, as described in the following.

Standard Similarity Measures and Cost Functions

We first recall some standard single and multimodal similarity measures described in the literature. They will be experimentally compared with the robust similarity measures introduced in this paper.

A classical similarity measure, widely used for the registration of single modal images is the least-squares cost function (Alpert *et al.*, 1996; Hajnal *et al.*, 1995). This similarity measure assumes that the two registered images differ only by an additive Gaussian noise (Christensen *et al.*, 1996), leading to the following least-squares cost function:

$$E(T_\Theta(x)) = \sum_x [I_{\text{ref}}(x) - I_{\text{reg}}(T_\Theta(x))]^2. \tag{3}$$

In the multimodal case, the interimage uniformity similarity measure (Woods *et al.*, 1993) has gained increasing popularity in recent years. This similarity measure assumes that a uniform region in the MR image corresponds to a uniform region in the PET image (uniformity hypothesis). In Wood's approach, the reference image is partitioned into its G gray level classes. The resulting spatial partition is simply projected onto the image to be registered, yielding the same segmentation of both images. For the image to be registered, the expected values μ_g, $g = 1, \ldots, G$ and the standard deviations σ_g, $g = 1, \ldots, G$ of each segmented region are computed. If the two images are correctly registered, the uniformity hypothesis implies that the ratio σ_g/μ_g is minimum over the entire volume (Woods *et al.*, 1993). The following image uniformity cost function is thus obtained

$$E(T_\Theta(x)) = \sum_{g=1}^{G} \frac{N_g}{N} \frac{\sigma_g(T_\Theta(x))}{\mu_g}, \tag{4}$$

where

$$\mu_g = \frac{1}{N_g} \sum_{x|I_{\text{ref}}(x)=g} I_{\text{reg}}(T_\Theta(x)), \tag{5}$$

and

$$\sigma_g(T_\Theta(x)) = \sqrt{\sum_{x|I_{\text{ref}}(x)=g} [I_{\text{reg}}(T_\Theta(x)) - \mu_g]^2}. \tag{6}$$

In Eq. (4), N represents the number of voxels in the images and N_g stands for the population of voxels having the value g in the reference image.

The image uniformity cost function has been devised for multimodal image registration. It may of course also be used for single modal image registration, although it is generally not as accurate as the least-squares method, as will be seen in our experiments.

Robust Similarity Measures

As already stated, the least-squares criterion (3) works at best under additive Gaussian noise assumptions (Christensen *et al.*, 1996). Its limits are now well known (Arun *et al.*, 1987; Umeyama, 1991). It is commonly accepted that least squares are sensitive to gross differences in images due to lesion evolution, to incomplete images, to non-Gaussian noise, or to "outliers." Outliers generally contribute too much to the overall solution since outlying points are assigned a high weight by the quadratic estimator (3). This remark also holds for the image uniformity cost function (4), which is based on standard image statistics (i.e., expected values and variances) and assumes a strict agreement between uniform regions in multimodal images. When a significant amount of outliers is present in the images to be registered, inaccurate registrations or even misregistrations may be obtained. To increase robustness, the cost function must thus be forgiving about outlying measurements.

To this end, several robust estimators have been introduced recently in computer vision. A review on robust estimators in computer vision may be found in Meer *et al.* (1990). A collection of different nonlinear robust estimators are presented by Black and Rangarajan (1996). A performance measure for a robust estimator is its breakdown point. The breakdown point is the largest fraction of data that can be arbitrarily bad and will not cause the solution to be arbitrarily bad.

The least median of squares regression (Rousseeuw, 1984), applied by Alexander and Somorjai (1996), is based on the minimization of the median of the squared residuals. The resulting estimator can resist to the effect of nearly 50% of contamination in the data. In the special case of simple regression, it corresponds to finding the narrowest strip covering half of the observations.

The MINPRAN algorithm (Stewart, 1995) also has a breakdown point of 50% and relies on random data distributions: random data sampling is used to search for the fit and the inliers to the fit that are least likely to have occurred randomly. These estimators have high breakdown points but also yield a high computational load, since they are based on random data sampling and sorting.

Another class of estimators, the M-estimators, which have attractive properties (i.e., satisfactory breakdown points and moderate computational cost), have been considered in computer vision (Black and Rangarajan, 1996). This class of robust estimators reduces the optimization problem to a simple, low-cost, weighted least-squares problem, as explained by Black and Rangarajan (1996) and Meer *et al.* (1990). They have a theoretical breakdown point of $1/(p+1)$, where p is the number of parameters to estimate (Meer *et al.*, 1990). However, in practice, it has been observed that for optical flow measurement, this family of robust estimators can tolerate roughly up to 35–45% of the data as outliers (Black and Jepson, 1996).

In our case, the robust M-estimation of parameters Θ is obtained by introducing a robust error norm ρ in the cost functions (3) and (4) (Meer *et al.*, 1990):

- For the single modality case, we define the robust least-squares cost function:

$$E(T_\Theta(x)) = \sum_x \rho[I_{\text{ref}}(x) - I_{\text{reg}}(T_\Theta(x)), C] \tag{7}$$

- For the multimodal case, we define the robust image uniformity cost function

$$E(T_\Theta(x)) = \sum_{g=1}^{G} \frac{N_g}{N} \widetilde{\sigma_g}(T_\Theta(x)), \tag{8}$$

where

$$\widetilde{\sigma_g}(T_\Theta(x)) = \sqrt{\sum_{x|I_{\text{ref}}(x)=g} \rho[I_{\text{reg}}(T_\Theta(x)) - \widetilde{\mu_g}, C]}, \tag{9}$$

and

$$\widetilde{\mu_g} = \arg\min_{\mu_g} \frac{1}{N_g} \sum_{x|I_{\text{ref}}(\bar{x})=g} \rho[I_{\text{reg}}(T_\Theta(x)) - \mu_g, C], \tag{10}$$

where C is a scale parameter and ρ is a nonquadratic error norm (penalty function) associated with the M-estimator.

Let us notice that the nonrobust cost functions (3) and (4) correspond to the special case $\rho(x, C) = x^2$. In the single modal case (7), the cost function is simply the robust error norm of the residual differences between

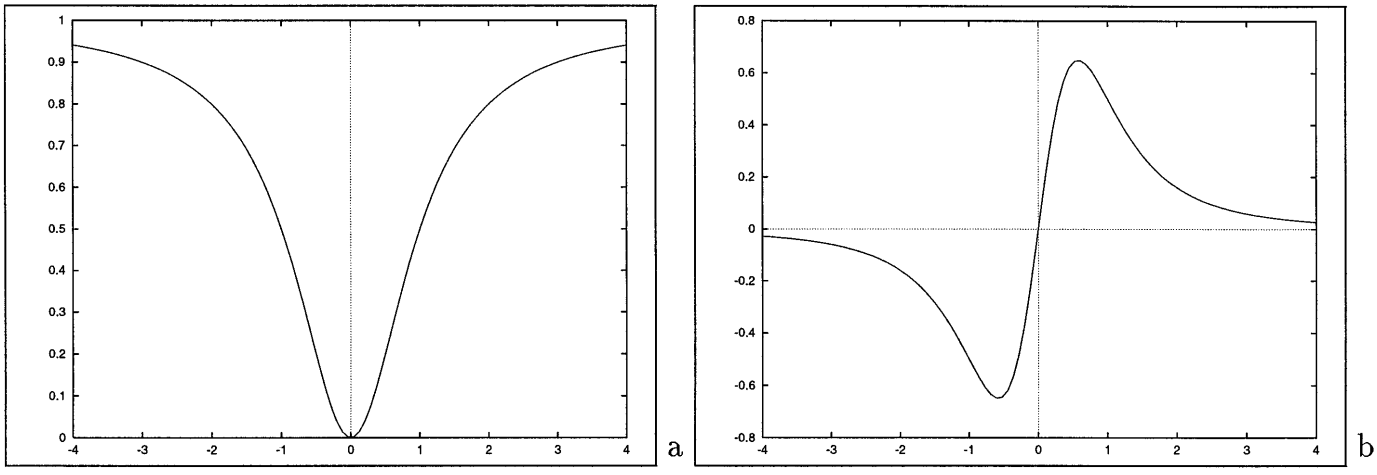

FIG. 1. The Geman–McClure robust estimator $\rho(x)$ (a) and its influence function (b).

the two registered images. In the multimodal case (8), a "robust variance" $\widetilde{\sigma_g}$ is computed for each region of the image to be registered, according to Eq. (9). This robust variance does not take into account outliers in the registered image, thanks to the robust error norm ρ. A robust estimation of the expected value $\widetilde{\mu_g}$ (10) of the region is simultaneously computed by the same M-estimator. Let us emphasize that *both* the expected value and the variance of each region must be computed using the robust M-estimator.

For the experiments presented in this paper, we take ρ to be the Geman–McClure estimator (Fig. 1)

$$\rho(x, C) = \frac{x^2}{C^2 + x^2},$$

which has successfully been applied to match temporal image sequences (Black and Jepson, 1996), for optical flow estimation (Black and Rangarajan, 1996), and for image restoration (Black and Rangarajan, 1996). Function ρ has a shape that rejects large residual errors. The influence function (Fig. 1) is the derivative of function $\rho(x)$ and characterizes the influence of the residuals. As can be seen in Fig. 1, as the magnitude of the residuals increases and grows beyond a point, its influence on the solution begins to decrease and the value of $\rho(x)$ approaches a constant. The scaling parameter C affects the point at which the influence of outliers begins to decrease. For the error norm used in our experiments, points x for which

$$|I_{\text{ref}}(T_\Theta(x)) - I_{\text{reg}}(x)| \geq \frac{C}{\sqrt{3}}$$

can be viewed as outliers, as the outliers rejection begins where $(\partial^2\rho)/(\partial x^2) = 0$. We have also experimented with the Tukey "biweight" estimator as well as with the truncated least-squares robust function (Black and Rangarajan, 1996). We applied the Geman–McClure estimator because it requires fewer calculations than the Tukey biweight and is more efficient than the truncated least squares. Furthermore, our purpose is not to find the best robust estimator for medical image registration but to underscore the fact that matching is more accurate when such a technique is applied.

The calculation of the registration parameters Θ involves the minimization of the nonlinear cost functions (7) or (8), which depend on the scale parameter C. We begin the optimization procedure with a high value for C. The value of C decreases during the minimization process following the formula $C = \alpha \cdot C$ with $0.8 < \alpha < 1$ until C reaches a predefined value. The effect of this procedure is that initially no data are rejected as outliers and a first, crude solution is obtained. During the following optimization steps, the influence of the outliers is gradually reduced by decreasing C, leading

TABLE 1

Single Modal Registration (2D)

Approach	Δt_x	Δt_y	$\Delta\hat{\theta}$
LS	0.77 ± 0.44	0.99 ± 1.09	3.90 ± 2.87
IU	2.12 ± 3.25	2.54 ± 1.91	4.34 ± 2.25
MI	0.29 ± 0.22	0.19 ± 0.12	0.15 ± 0.14
RLS	0.27 ± 0.26	0.20 ± 0.07	0.14 ± 0.77
RIU	0.18 ± 0.12	0.29 ± 0.23	0.66 ± 0.67

Note. A set of 2D slices was artificially transformed using 20 different rigid transformations and the images were corrupted at 25% by salt-and-pepper noise. The average and the standard deviation of the registration errors computed from the 20 registrations are presented for the different approaches. The translation error is given in pixels and the rotation error in degrees.

TABLE 2

Single Modal Registration (3D)

Approach	Δt_x	Δt_y	Δt_z	$\Delta\hat{\theta}_x$	$\Delta\hat{\theta}_y$	$\Delta\hat{\theta}_z$
LS	2.30 ± 1.75	2.53 ± 1.56	2.77 ± 1.83	4.71 ± 2.89	5.33 ± 3.40	5.05 ± 3.51
IU	1.49 ± 1.40	1.56 ± 1.41	1.93 ± 1.63	3.75 ± 2.03	3.65 ± 2.54	2.99 ± 3.06
MI	0.05 ± 0.06	0.22 ± 0.15	0.09 ± 0.14	0.35 ± 0.35	0.27 ± 0.32	0.44 ± 0.69
RLS	0.04 ± 0.07	0.16 ± 0.11	0.06 ± 0.10	0.41 ± 0.21	0.16 ± 0.22	0.33 ± 0.24
RIU	0.09 ± 0.05	0.18 ± 0.14	0.10 ± 0.05	0.22 ± 0.34	0.24 ± 0.17	0.40 ± 0.59

Note. A set of 3D image volumes was artificially transformed using 20 different rigid transformations and the images were corrupted at 25% by salt-and-pepper noise. The average and the standard deviation of the registration errors computed from the 20 registrations are presented for the different approaches. The translation error is given in pixels and the rotation error in degrees.

to a reliable estimation of the rigid transformation parameters, which is robust to gross image differences. In other experiments we have computed C as the noise variance computed online during the initial segmentation procedure. The two strategies provide us with approximately the same results.

Fast Stochastic Optimization

The robust and nonrobust cost functions considered previously all lead to highly nonlinear estimation problems, involving many local minima. This is the case for all standard similarity functions.

As a consequence, classical continuous or discrete deterministic optimization algorithms are trapped by local minima unless a specific additional procedure is introduced to handle large initial misregistrations. For instance Studholme *et al.* (1997) report successful matching with initial misregistrations up to 30 mm and 30°. These results are obtained by taking a starting transformation estimate that is within the capture range of the registration measure. Other standard approaches include randomly distributed initial configurations (Studholme *et al.*, 1997) or multiresolution data decompositions (Alexander and Somorjai, 1996).

An alternative approach (adopted here) is to consider stochastic optimization algorithms that theoretically ensure convergence toward the global minimum. Stochastic optimization methods are thus able to handle large misregistrations without any *a priori* knowledge of the characteristics of the similarity function.

The fast stochastic optimization process, which is applied here generally, yields close to the optimal solutions. The optimization technique used in our implementation relies on an iterative fast simulated annealing algorithm (Kirkpatrick *et al.*, 1983; Aarts and Laarhoven, 1985) based on the Gibbs sampler dynamics (Geman and Geman, 1984). A high value is adopted for the initial temperature in the simulated annealing procedure and a fast exponentially decreasing temperature schedule is considered instead of the optimal logarithmic descent (readers not familiar with simulated annealing are referred to Aarts and Laarhoven (1985)).

The solution obtained after a given number of steps is further improved by a deterministic extension of the above algorithm, known as iterated conditional modes (ICM; Besag, 1986). ICM is a simulated annealing technique with the temperature variable set to zero. Only configurations decreasing the cost function are accepted. It has fast convergence properties and local minima are not a problem, since the first stochastic optimization step provides a good initialization. To speed up the algorithm, multigrid data processing is implemented, as explained previously.

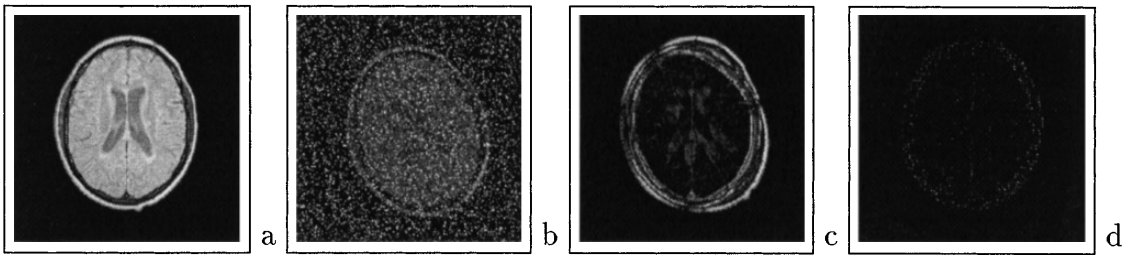

FIG. 2. 2D robust registration. (a) Reference image. (b) Image in (a) rotated by 20°, translated by 10 pixels along the x axis and 10 pixels along the y axis and corrupted at 25% with salt-and-pepper noise. (c) Difference between the noise-free registered image and the image in (a) when the nonrobust technique is applied. (d) Difference between the noise-free registered image and the image in (a) using the Geman–McClure robust estimation function.

TABLE 3

Single Modal Test Object Registration (2D)

Approach	Δt_x	Δt_y	$\Delta\hat{\theta}$
LS	0.29 ± 0.22	0.18 ± 0.18	0.26 ± 0.25
IU	0.43 ± 0.51	0.45 ± 0.88	0.75 ± 0.71
MI	0.25 ± 0.15	0.22 ± 0.29	0.29 ± 0.30
RLS	0.23 ± 0.20	0.19 ± 0.21	0.17 ± 0.23
RIU	0.33 ± 0.24	0.27 ± 0.29	0.27 ± 0.27

Note. A 2D test object was acquired with 35 different rigid transformations. The average and the standard deviation of the registration errors are presented for the different approaches. The translation error is given in pixels and the rotation error in degrees.

Interpolation

A large number of interpolations are involved in the registration process. The accuracy of the rotation and translation parameter estimates is directly related to the accuracy of the underlying interpolation model. Simple approaches such as the nearest neighbor interpolation are commonly used because they are fast and simple to implement, though they produce images with noticeable artifacts. More satisfactory results can be obtained by small-kernel cubic convolution techniques, trilinear interpolation, bicubic spline interpolation, or convolution-based interpolation. According to the sampling theory, optimal results are obtained using sinus cardinal interpolation, but at the expense of a high computational cost. As a compromise, we have used a fast nearest neighbor interpolation technique in the first optimization steps. At the end of the algorithm, the registration parameters are refined using a trilinear (3D) interpolation that preserves the quality of the image to be registered. This technique has been revealed to be fast and efficient.

Data Acquisition

To evaluate the different cost functions and registration algorithms, the following data sets have been acquired:

- 2D MR images of a phantom and of selected multiple sclerosis patients were acquired on a 0.28-T Bruker system. Each MR image set was obtained with a multislice multiecho sequence (echo time TE = 15 ms, repetition time TR = 2000 ms). The size of the images was 256 × 256 (FOV = 25.6 cm) with a pixel size of 1 mm × 1 mm and a slice thickness of 5 mm.
- 3D MR scans were acquired on patients with partial complex epilepsy using a 2-T Bruker system. The images were obtained with a gradient echo sequence (TR = 40 ms and TE = 8.5 ms, flip angle = 45°), image size is 128 × 128 × 128 (FOV = 25.6 cm).
- SPECT imaging was performed on a double-headed camera (Elscint Helix) with low-energy and high-resolution parallel-hole collimators, using 700 M Bq of ^{99m}Tc − HMPAO or ^{99m}Tc − ECD. The camera was operated in the "stop and shoot" mode with acquisition at 3° intervals, acquiring 120 views at 40 s per interval (120 projections, 64 × 64 matrix). Slices were reconstructed with a matrix of 128 × 128. System resolution was measured at 8 mm full width at half-maximum (FWHM) in all planes at the center of the field.

For patients with partial complex epilepsy, the interictal SPECT studies were performed when patients had been seizure-free for at least 24 h. EEG recording was performed during isotope injections to ensure interictal status at the time of injection. For ictal studies, patients underwent continuous video-EEG monitoring and were injected during ongoing spontaneous seizure activity.

All studies involving human subjects were performed with the permission of the research committee.

To test the registration algorithms with ground truth data, part of the MR images were acquired with different offsets in demodulation frequency to simulate translations of the data set. Different directions of the read gradient were also used to generate rotations. With the above manipulation, the true values of the 3D translation and rotation parameters were known accurately for these data sets and could be used to compare the performances of the different approaches.

Both computation and display were performed on a Hewlett-Packard 715/80 workstation by using 2D–

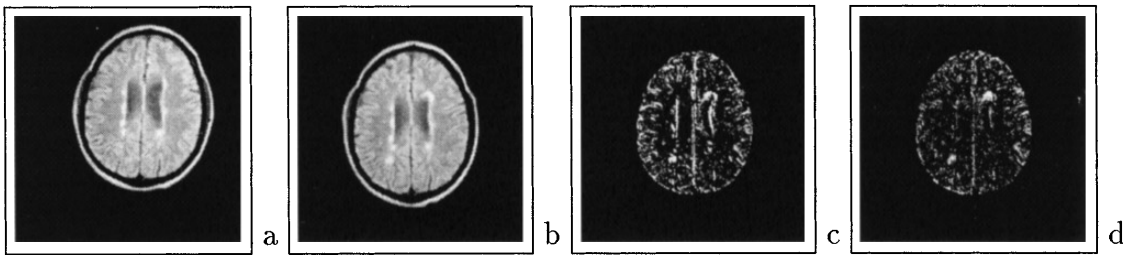

FIG. 3. (a) A multiple sclerosis patient's MR image. (b) Image of the same patient acquired some months later. (c) Difference between the registered image and the image in (b) when the least-squares technique is applied. (d) Difference between the registered image and the image in (b) using the Geman–McClure robust estimation function.

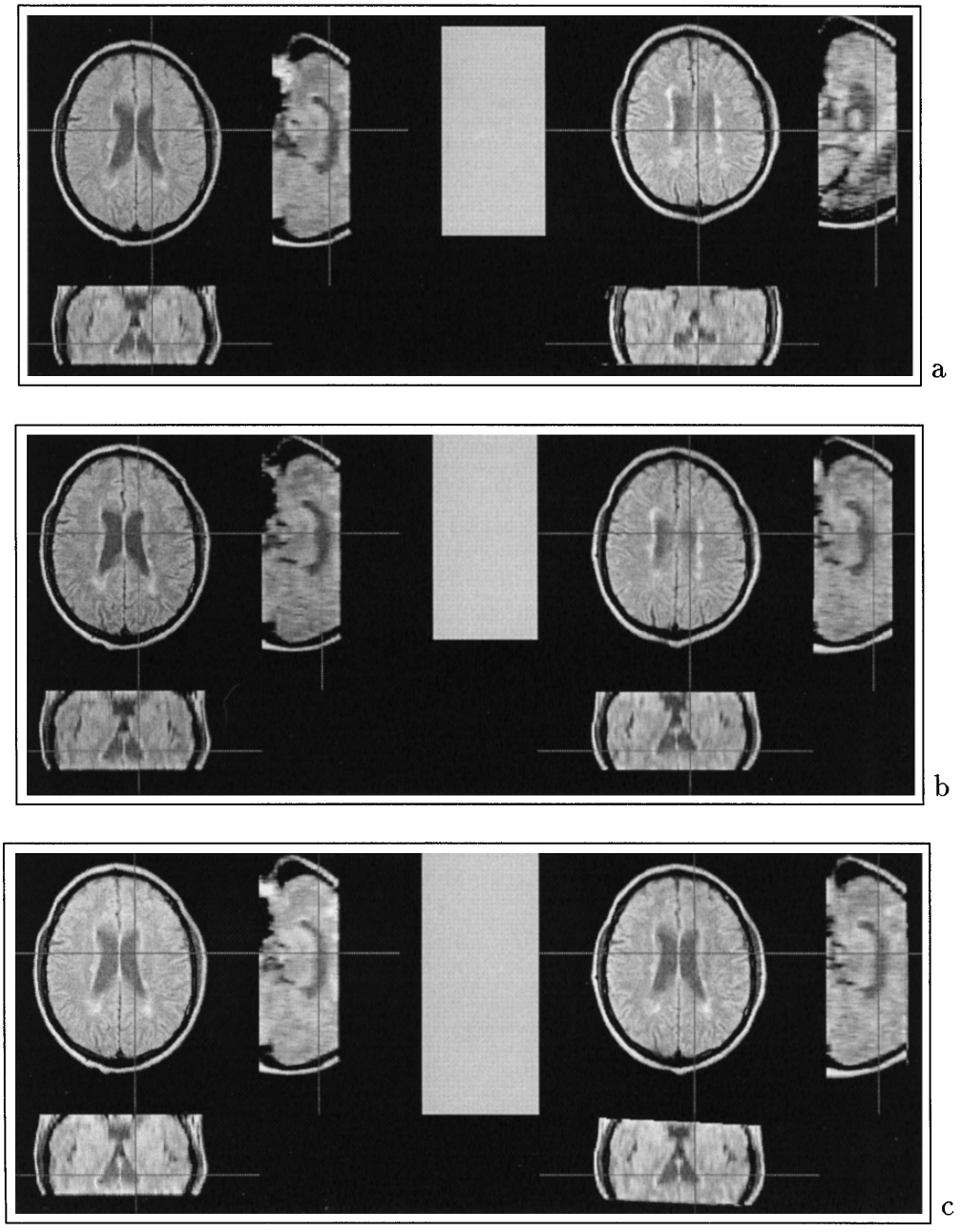

FIG. 4. MRI/MRI registration. The two MR volumes of a patient presenting multiple sclerosis are presented (a) before registration and (b) after registration by the LS technique. Notice that the horizontal slices do not correspond. (c) After robust registration by the RLS technique. The horizontal slices show better correspondence.

3D image analysis software (MEDIMAX) developed at the IPB. This software, running under Unix, was developed in C language and uses the standard graphics interface X11/R5 and the Motif windows manager. All registrations techniques presented in this paper were implemented under this software environment and are easily available to users. The software is presented on the laboratory's web server (http://alsace.u-strasbg.fr).

RESULTS

We have compared the robust least-squares (RLS) (7) and robust image uniformity (RIU) (8) approaches to the standard least-squares (LS) method (3) and to the image uniformity (IU) technique (4). Another recently published approach, based on entropy measures and on the maximization of the mutual information (MI) of the two images to be registered (Wells *et al.*, 1996), has also been implemented for comparison purposes.

As mentioned by several authors (Maes *et al.*, 1997; Studholme *et al.*, 1997), the mutual information criterion works relatively well even if no brain segmentation is applied. However, it is accepted that only misregistrations near the optimal solution are recovered by this strategy (Maes *et al.*, 1997) or the gold standard registration accepts larger errors (Studholme *et al.*, 1997). For multimodal registration (MR/SPECT), nonbrain structures have been removed as explained under Materials and Methods. The above manipulation was applied in order to compare the similarity metrics on the same basis and to ensure subvoxel accuracy. Several tests have been also performed without removing nonbrain structures.

Moreover, the improved methods for image registration (Eq. (1)) proposed by Alpert *et al.* (1996) are applied for both standard and robust metrics in order to handle intensity mismatch. By these means, all the similarity metrics compared in our study were tested under the same conditions.

Single Modal Registration

Registration experiments were performed with both 2D and 3D images. A first class of experiments consisted of applying a known transformation (translations and rotations) to a set of MRI slices or volumes to create a second image set. Twenty-five percent of the transformed images were then corrupted by salt-and-pepper noise, to simulate outliers. For each method, the estimated registration parameters were compared to the true parameters to determine the accuracy of the registration. Statistics on the registration errors were computed on a set of 20 different registrations problems, involving translation parameters between -20 and $+20$ voxels and rotations between $-30°$ and $+30°$. As we can see in Table 1 (2D images) and Table 2 (3D images), the robust algorithms achieved subvoxel registration errors while the nonrobust (LS and IU) techniques failed. The MI method also achieved subvoxel registration but its performance is slightly inferior to the results obtained by the RLS technique. Figure 2c shows an example where the standard method (LS) failed to correctly register the MR slices shown in Figs. 2a and 2b, but where the RLS achieved accurate matching by discarding the outliers. The registration error shown in Figs. 2c and 2d is the residual squared image difference after registration. The registration errors in Figs. 2c and 2d are normalized to the maximum display value for better visualization.

Complementary experiments, with known ground truth, were obtained with a 2D test object, acquired under 35 different rigid transformations by modifying the read and phase gradients during acquisition, as explained previously. Table 3 presents the registration errors for the different techniques in this case. In the absence of significant noise, all of the techniques achieved subvoxel accuracy but the RLS gave the best results and appears to be a good choice for the single modal registration problem. Let us notice that the image uniformity approaches (IU or RIU) are not appropriate methods for single modal image registration, as can be seen from the results in Table 1, 2, and 3 (as already noticed, they have rather been devised for multimodal images).

In addition, we have applied the RLS algorithm to a set of MRI slices of a multiple sclerosis patient, acquired at different dates (Rumbach *et al.*, 1996). Figure 3 illustrates an accurate alignment where small differences due to lesion evolution, which were not well distinguished previously due to misalignment by standard methods (Fig. 3c), are now identified by simple subtraction (see Fig. 3d). Let us notice that the image in Fig. 3d illustrates the absolute value of the difference between the reference image and the registered image. By these means, both the straight and the inverse subtractions are represented. The robust algorithm achieved better registration than the standard algorithm (Fig. 3). Fewer registration artifacts are observed on the cortical sulcus, the falx, and the periventricular hyperintensities, and multiple sclerosis lesion evolution is clearly displayed. The robust registration technique allowed better follow-up of the disease.

Finally, a 3D MR/MR registration example where the robust least-squares metric achieved better registration is presented (Fig. 4). In this case, both lesion evolution and incomplete acquisition are handled by the robust function. The MR volumes are acquired by a multislice sequence. Therefore, the two volumes only

TABLE 4

Multimodal (MRI/SPECT) Registration (3D)

Approach	Δt_x	Δt_y	Δt_z	$\Delta\hat{\theta}_x$	$\Delta\hat{\theta}_y$	$\Delta\hat{\theta}_z$
IU	3.85 ± 5.59	3.02 ± 4.78	4.16 ± 4.38	8.33 ± 4.51	6.23 ± 3.52	6.80 ± 4.15
MI	1.41 ± 0.74	1.38 ± 1.23	2.06 ± 1.29	0.94 ± 1.58	1.04 ± 1.15	1.36 ± 0.77
RIU	0.82 ± 0.53	0.61 ± 0.50	0.83 ± 0.60	0.21 ± 0.48	1.14 ± 0.26	0.71 ± 0.94

Note. A set of 3D SPECT image volumes manually preregistered by an expert to its MRI counterpart was artificially transformed using 20 different translation and rotation parameters and corrupted at 25% by salt-and-pepper noise. The average and the standard deviation of the registration errors are presented for the different approaches. The translation error is given in pixels and the rotation error in degrees.

partially overlap. The least-squares cost function handled displacement along the z axis but it is clearly illustrated that horizontal slices do not correspond (Fig. 4b). On the other hand, the robust version of the least-squares measure achieved better registration (Fig. 4c).

Multimodal Registration

To evaluate the multimodal image registration algorithms, a 3D SPECT image volume was manually registered to its corresponding MRI volume with the aid of a neurologist. The manually registered SPECT image was then transformed using the same 3D translation and rotation parameters as for the previously described experiments. To simulate outliers, 25% of the SPECT image was corrupted by salt-and-pepper noise. The robust image uniformity technique RIU has been compared to the image uniformity IU technique and to the MI method (Wells *et al.*, 1996), which is also suited

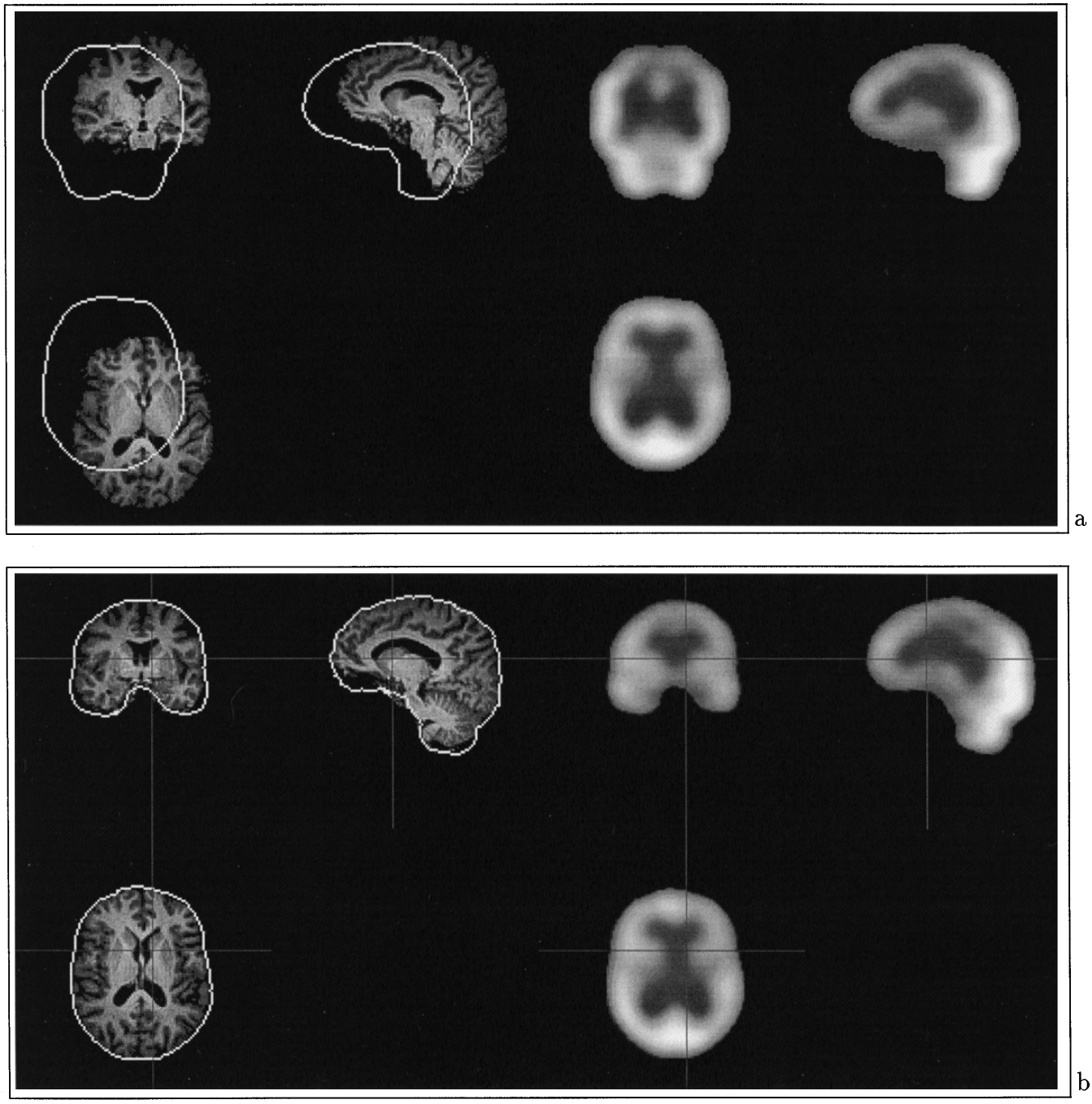

FIG. 5. Robust MRI/SPECT registration. The SPECT and MRI volume with the SPECT contours superimposed are shown (multiplanar visualization). (a) Before registration. (b) After robust registration by the RIU technique.

to multimodal image registration. Table 4 illustrates the robustness of our technique to outliers. The error for the RIU method is around 1 pixel for the translation and 1° for the rotation. This is significantly more accurate than the IU approach. We also notice the good performance of the MI technique, which provides results that are always better than the IU but generally slightly inferior to RIU.

Figure 5 shows a real example of a patient's SPECT image volume (ictal and interictal) registered with respect to its MRI counterpart by the robust algorithm. From the same SPECT image volume, the hyperinten-

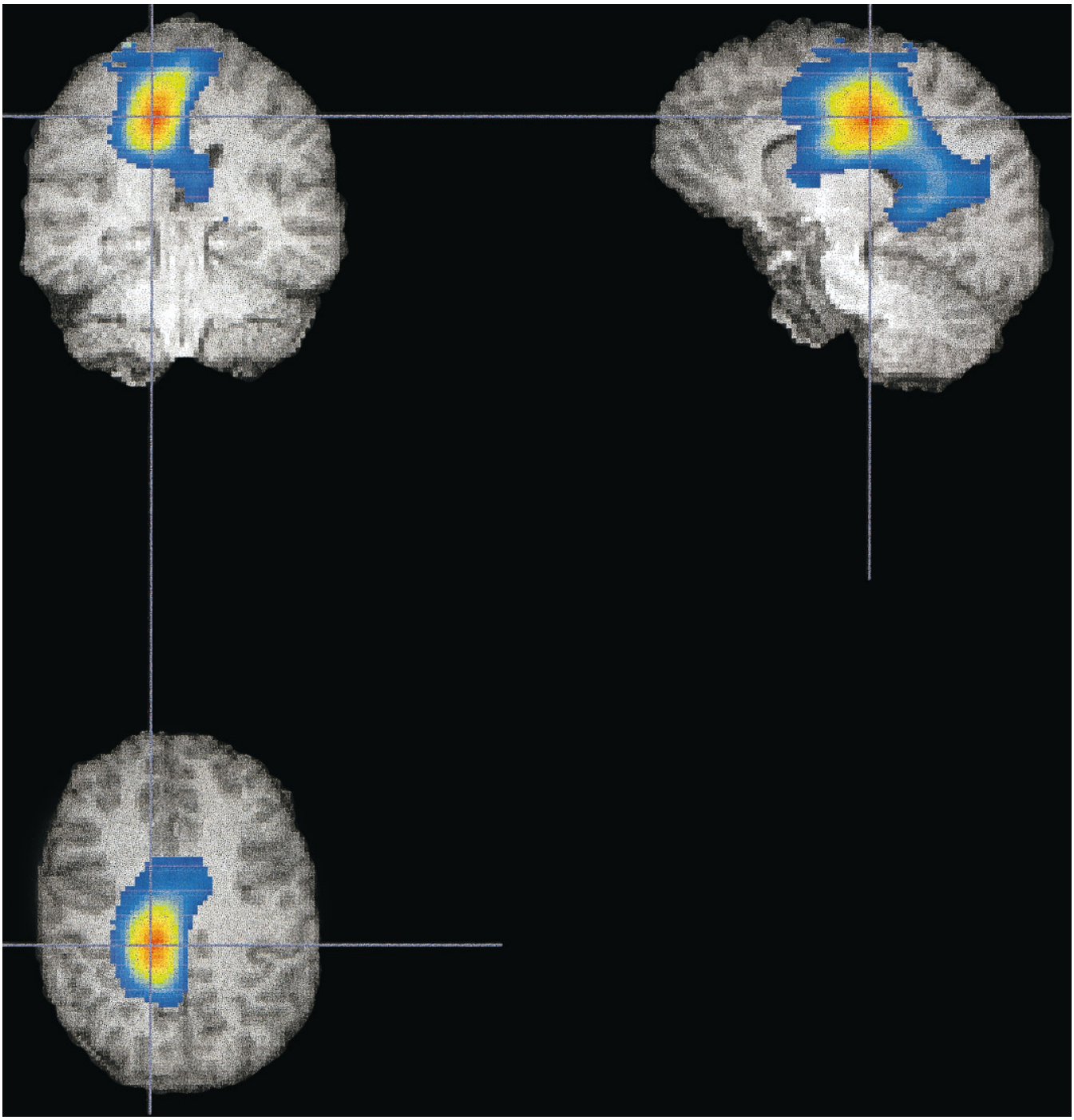

FIG. 6. 3D MRI/SPECT representation of a patient presenting partial complex seizures of the right temporal origin: the difference image (ROI) between ictal SPECT and interictal SPECT, demonstrating areas of increased perfusion, is shown superimposed onto the corresponding MR image.

sity region has been segmented and superimposed onto the MR image (Fig. 6).

Robustness to the presence of nonbrain structures has also been examined. All of the similarity metrics (IU, MI, RIU) are applied to MRI/SPECT registration without prior removal of skull and scalp from the MR volumes. The shortcoming of initializing close to the optimal solution is overcome by the stochastic optimization technique for all metrics. Our first results confirm that the IU criterion is trapped by many local minima (Alexander and Somorjai, 1996) as it considers outliers (nonbrain structures) as a part of the data (Fig. 7a). On the other hand, mutual information (Fig. 7b) and M-estimators provide (Fig. 7c) approximately similar results with relatively good precision. However, their accuracies are inferior to those obtained when brain structures are segmented (due to the presence of large outliers).

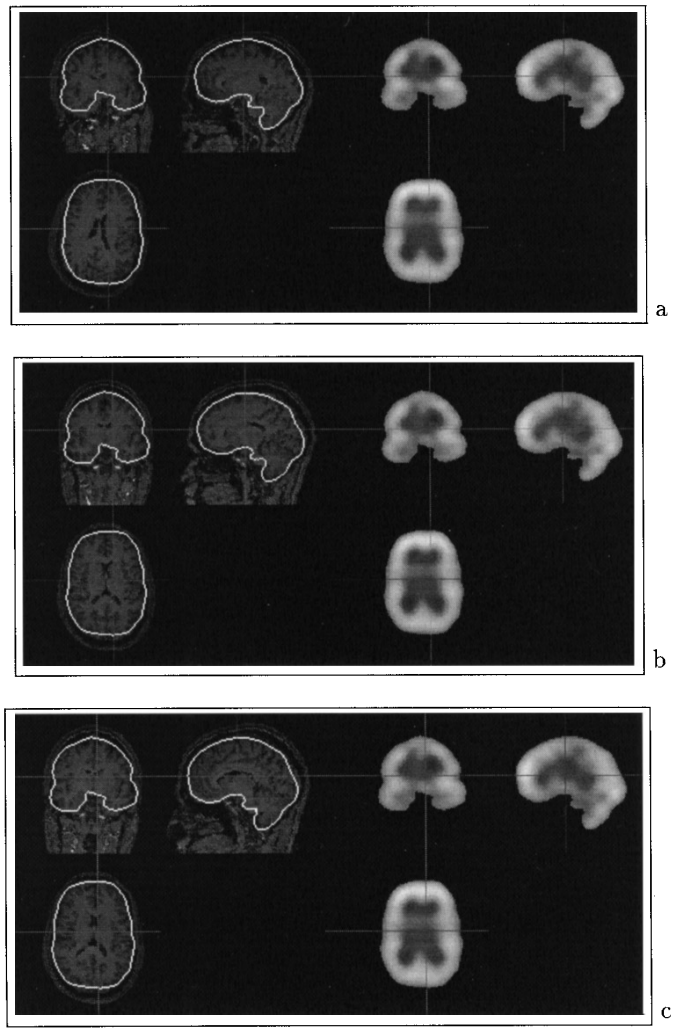

FIG. 7. MRI/SPECT registration without removing nonbrain structures. (a) IU metric. (b) MI metric. (c) RIU metric.

DISCUSSION AND CONCLUSION

The registration methods described in this paper were motivated by the algorithm proposed by Woods *et al.* (1992) and Hajnal *et al.* (1995) in the case of single modality medical image registration and by the model proposed in Woods *et al.* (1993) in the case of multimodal image matching. These approaches have been improved by the nonstraightforward extensions proposed in this paper. The new robust multigrid stochastic registration technique has two major advantages over previous methods:

- No manual initialization near the optimal solution is required to obtain an accurate registration. Local minima, a major problem in standard medical image registration techniques, are avoided automatically by the use of fast simulated annealing optimization algorithms. This results in a fully data-driven method that requires no human interaction.
- Gross image differences are taken into account efficiently by robust estimation techniques. Until now, dissimilarities (due to lesion evolution, etc.) were not taken into account. The robust functions implemented in our algorithm determine whether or not a measure is an outlier. To our knowledge, robust registration has never been evoked for multimodal images until now.

We have compared our approach to the maximization of the mutual information technique (Wells *et al.*, 1996) and to the commonly used image uniformity algorithm (Woods *et al.*, 1992, 1993). The IU algorithm does not perform well when the images exhibit significant differences, since its cost function, based on standard image statistics, does not account for outliers. The MI method presents a good robustness to outliers but its performance is not as good as that observed with the robust image uniformity technique.

In addition, the robust similarity metrics proposed here take implicitly into account images that partially overlap (Fig. 4). In general, standard techniques handle incomplete acquisitions by ignoring data falling outside the field of view of the floating image. The above manipulation is efficient in cases where one of the images is a subset of the other image of the pair, but fail when each image carries information not represented in the other image. Robust techniques intrinsically reject nonoverlaping structures (Black and Rangarajan, 1996).

Furthermore, as it is shown by the test-object results, in the absence of outliers both standard and robust methods provide subvoxel accuracies. Our concern was to account for critical situations where standard methods fail (nonoverlapping data, lesion evolution, preoperative and postoperative images, etc.).

Finally, let us notice that the LS and RLS techniques require approximately the same computation times (20

min cpu time for $128 \times 128 \times 128$ images on our HP 715/80 workstation). On the same data set, the IU method takes 35 min, the MI technique takes 40 min, and the RIU method needs 1 h of cpu time. As can be seen, the additional computational complexity introduced by the robust estimation remains moderate and these methods may thus be used with profit to improve the accuracy in many critical multimodal image registration problems.

ACKNOWLEDGMENTS

This study was supported by the Commission of the European Communities, DG XII, in the framework of the TMR program (Training and Mobility of Researchers), Contract No. ERBFMIBCT960701, and by the "Groupement d'Intérêt Scientifique" (CNRS, CEA, INRIA, MENESR) "Sciences de la Cognition."

REFERENCES

Aarts, E. H. L., and Laarhoven, P. J. M. V. 1985. Statistical cooling: A general approach to combinatorial optimization problems. *Philips J. Res.* **40**(4):193–226.

Alexander, M., and Somorjai, R. L. 1996. The registration of MR images using multiscale robust methods. *Magn. Reson. Imaging* **14**(5):453–468.

Alpert, N. M., Berdichevsky, D., Levin, Z., Morris, E. D., and Fischman, A. J. 1996. Improved methods for image registration. *NeuroImage* **3:**10–17.

Alpert, N. M., Bradshaw, J. F., Kennedy, D., and Correia, J. A. 1990. The principal axis transformation: A method for image registration. *J. Nucl. Med.* **31:**1717–1722.

Arata, L. K., Dhawan, A. P., Broderik, J. P., Gaskil-Shipley, M., Levy, A. V., and Volkow, N. D. 1995. Three-dimensional anatomical model-based segmentation of MR brain images through principal axis registration. *IEEE Trans. Biomed. Eng.* **42**(1):1069–1078.

Arun, K. S., Huang, T. S., and Blostein, S. D. 1987. Least squares fitting of two 3D point sets. *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(3):698–700.

Bajcsy, R., and Kovačič, S. 1989. Multiresolution elastic matching. *Comput. Vis. Graphics Image Proc.* **46:**1–21.

Besag, J. 1986. On the statistical analysis of dirty pictures. *J. R. Statist. Soc.* **48**(3):259–302.

Black, M. J., and Jepson, A. D. 1996. Eigen tracking: Robust matching and tracking of articulated objects using a view-based representation. In *Proceedings of the European Conference on Computer Vision,* pp. 1–14, Cambridge, UK.

Black, M. J., and Rangarajan, A. 1996. On the unification of line processes, outliers rejection and robust statistics in early vision. *Int. J. Comput. Vis.* **19**(1):57–91.

Brown, L. G. 1992. A survey of image registration techniques. *ACM Comput. Surv.* **24**(4):325–376.

Christensen, G., Miller, M., Vannier, M., and Grenander, U. 1996. Individualizing neuro-anatomical atlases using a massively parallel computer. *IEEE Comput.*, 32–38.

Davatzikos, C. 1996. Spatial normalization of 3D brain images using deformable models. *J. Comput. Assist. Tomogr.* **20**(4):656–665.

Dhawan, A. P., Arata, L. K., Levy, A. V., and Mantil, J. 1995. Iterative principal axes registration method for analysis of MR-PET images. *IEEE Trans. Biomed. Eng.* **42**(11):1079–1087.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., and Frackowiak, R. S. J. 1995. Spatial registration and normalization of images. *Hum. Brain Mapping* **2:**165–189.

Geman, S., and Geman, D. 1984. Stochastic relaxation, Gibbs distribution and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(6):721–741.

Hajnal, J., Saeed, N., Soar, E. J., Oatridge, A., Young, I. R., and Blyder, G. M. 1995. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *J. Comput. Assist. Tomogr.* **19**(2):289–296.

Heitz, F., Perez, P., and Bouthemy, P. 1994. Multiscale minimization of global energy functions in some visual recovery problems. *Comput. Vis. Graphics Image Proc. Image Understanding* **59**(1):125–134.

Herbin, M., Venot, A., Devaux, J. Y., Walter, E., Lebruchec, F., Dubertet, L., and Roucayrol, J. C. 1989. Automated registration of dissimilar images: Application to medical imagery. *Comput. Vis. Graphics Image Proc.* **47:**77–88.

Kirkpatrick, S., Gellat, C. D., and Vecchi, M. P. 1983. Optimization by simulated annealing. *Science* **200:**671–680.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. 1997. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging* **16**(2):187–198.

Meer, P., Mintz, D., Rosenfeld, A., and Kim, D. Y. 1990. Robust regression methods for computer vision: A review. *Int. J. Comput. Vis.* **6**(1):59–70.

Minoshima, S., Berger, K., Lee, K., and Mintum, M. 1992. An automated method for rotational correction and centering of three-dimensional functional brain images. *J. Nucl. Med.* **33:**1579–1585.

Otsu, N. 1979. A threshold selection method from grey-level histograms. *IEEE Trans. Systems Man Cybernetics* **9**(1):62–66.

Pelizzari, C. A., Chen, G. T. Y., Spelbring, D. R., Weischelbaum, R. R., and Chen, C. T. 1989. Accurate three-dimensional registration of CT, PET and/or MR images of the brain. *J. Comput. Assist. Tomogr.* **13**(1):20–26.

Pietrzyk, U., Herholz, K., Jacobs, A., Mielke, R., Slansky, I., Wurker, M., and Heiss, W. D. 1994. An interactive technique for three-dimensional image registration: Validation for PET, SPECT, MRI and CT brain images. *J. Comput. Assist. Tomogr.* **35:**2011–2018.

Rousseeuw, P. J. 1984. Least median of squares regression. *J. Am. Statist. Assoc.* **79:**871–880.

Rumbach, L., Racadot, E., Armspach, J. P., Namer, I. J., Bonneville, J. F., Wijdenes, J., Marescaux, C., Herve, P., and Chambron, J. 1996. Biological assessment and MRI monitoring of the therapeutic efficacy of a monoclonal anti-T CD4 antibody in multiple sclerosis patients. *Multiple Sclerosis,* 207–212.

Siomka, P. J., Hurwitz, G. A., Stephenson, J., and Cradduck, T. 1995. Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm. *J. Nucl. Med.* **36:**1115–1122.

Stewart, C. V. 1995. MINPRAN: A new robust estimator for computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(10):925–938.

Studholme, C., Hill, D., and Hawkes, D. 1997. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med. Phys.* **24**(1):25–35.

Turkington, T., Hoffman, J., Jaszczak, R., MacFall, J. R., Harris, C. C., Kilts, C. D., Pelizzari, C., and Coleman, R. E. 1995. Accuracy of surface fit registration for PET and MR brain images using full and incomplete brain surfaces. *J. Comput. Assist. Tomogr.* **19**(1): 117–124.

Umeyama, S. 1991. Least-squares estimation of transformation parameters between two point patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(4):376–380.

Van den Elsen, P., Maintz, J. B. A., Pol, E. J. D., and Viergever, M. A. 1995. Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Trans. Med. Imaging* **14**(2):384–396.

Van den Elsen, P., Paul, E. J. D., and Viergever, M. A. 1993. Medical image matching—A review with classification. *IEEE Eng. Med. Biol.* **12**(1):26–39.

Wells, W., III, Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R. 1996. Multimodal volume registration by maximization of mutual information. *Med. Image Anal.* **1**(1):33–51.

Woods, R. P., Cherry, S. R., and Mazziota, J. C. 1992. Rapid automated algorithm for aligning and reslicing PET images. *J. Comput. Assist. Tomogr.* **16**(4):620–633.

Woods, R. P., Mazziota, J. C., and Cherry, S. R. 1993. MRI-PET registration with automated algorithm. *J. Comput. Assist. Tomogr.* **17**(4):536–546.