

Single and multimodal subvoxel registration of dissimilar medical images using robust similarity measures

Christophoros NIKOU^{†,‡} Fabrice HEITZ[†] Jean-Paul ARMSPACH[‡]
Izzie-Jacques NAMER[‡]

[†]Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection,
Université Louis Pasteur de Strasbourg
4 Bd. Sébastien Brant, 67400 Illkirch, France

[‡]Institut de Physique Biologique, Faculté de Médecine,
Université Louis Pasteur de Strasbourg
4 rue Kirschleger, 67085 Strasbourg, France

ABSTRACT

Although a large variety of image registration methods have been described in the literature, only a few approaches have attempted to address the rigid registration of medical images showing gross dissimilarities (due for instance to lesion evolution). In the present paper, we develop data driven registration algorithms, relying on robust pixel similarity metrics, that enable an accurate (subvoxel) rigid registration of dissimilar single or multimodal 2D/3D images. In the proposed approach, gross dissimilarities are handled by considering similarity measures related to robust M-estimators. A “soft re-descending” estimator (the Geman-McClure ρ -function) has been adopted to reject gross image dissimilarities during the registration. The registration parameters are estimated using a top down stochastic multigrid relaxation algorithm. Thanks to the stochastic multigrid strategy, the registration is not affected by local minima in the objective function and a manual initialization near the optimal solution is not necessary. The proposed robust similarity metrics compare favourably to the most popular standard similarity metrics, on patient image pairs showing gross dissimilarities. Two case studies are considered : the registration of MR/MR and MR/SPECT image volumes of patients suffering from multiple sclerosis and epilepsy.

Keywords: multimodal image registration, stochastic optimization, dissimilar images, robust estimation, Magnetic Resonance Imaging (MRI), Single Photon Emission Computed Tomography (SPECT).

1. 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 last 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.¹ “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 both for single and multimodal medical image registration. In the case of single modal image registration, the cost function is generally

Other author information:

Send correspondence to C.N. LSIIT-GRI, 4 Bd S. Brant, 67400 Illkirch, France

C.N.: Email: nikou@mondrian.u-strasbg.fr; Telephone: ++ 33 3 88 65 50 51; Fax ++ 33 3 88 65 52 49

F.H. Email: Fabrice.Heitz@ensps.u-strasbg.fr; Telephone: ++ 33 3 88 65 50 35; Fax: ++ 33 3 88 65 52 49

J.P.A. Email: armspach@alsace.u-strasbg.fr; Telephone: ++ 33 3 88 14 48 44; Fax: ++ 33 33 88 37 14 77

I.J.N. Email: namer@alsace.u-strasbg.fr; Telephone: ++ 33 3 88 14 48 40 Fax: ++ 33 3 88 37 14 77

related to a Gaussian sensor model² or, equivalently, to least squares estimation expressed through different cost functions.³⁻⁷ Other similarity measures, based on standard image statistics (mean, variance⁸ or entropy measures⁹) have been devised in the case of multimodal medical images.

Similarity measure-based approaches have been shown to be very efficient for coping with standard 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 non linear, yielding many local minima in the cost function. As a consequence, registration algorithms have to be initialized close to the optimal (or desired) solution in order to get satisfactory solutions. A second important limitation of existing algorithms is related to the underlying sensor models. For the single modal case, the Gaussian sensor model assumes that the two images to be aligned differ only by additive Gaussian noise.² 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 by Woods *et al.* for MR/PET image registration.⁸ Woods' 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. Besides, due to Compton scattering in SPECT images, anatomical structures do not occupy the same volume in MRI and SPECT image sequences. Single or multimodal image pairs may thus differ significantly, and these differences are not uniform. The standard similarity-based approaches do not model these information differences and as a consequence, are not robust with respect to them. Standard methods may thus result in inaccurate registrations or even mis-registrations.

In this paper we propose a similarity measure-based approach that addresses the above mentioned shortcomings. Robust statistics are introduced both in single modal and multimodal image registration similarity measures in order to account for significant image differences (outliers). The sensitivity of the registration algorithm to local minima in the similarity measure is also reduced by using a fast multigrid 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.

Cost functions for the registration of *2D single modal images*, that are to a certain extent robust to image changes have already been described.¹⁰ Herbin *et al.* make use of deterministic and stochastic sign change criteria to provide robust registrations in critical situations corresponding for instance to lesion evolution.¹⁰ Another robust statistics-based approach, has been proposed recently and independently¹¹ for the registration of *2D single modality images*. The above mentioned approach¹¹ is feature-based and relies on a least median of squares robust estimator.¹² Contrary to the method described below, its application to 3D multimodal images has not been described and is not straightforward.

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 favourably with the other standard similarity measure-based approaches, such as the least squares cost function,⁴ the image uniformity cost function⁸ and the entropy-based similarity measure.⁹ Sub-voxel accuracies are obtained in all cases by the robust methods. Experimental results on 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 partial epileptic patients onto the MRI of the same patients.

2. 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 sequence of multiresolution grids, using a top-down approach starting from the coarsest resolution level.¹³ The solution obtained at a given resolution level is interpolated and forwarded to the next, finer resolution. The algorithm first carries out the calculations for every 81^{st} voxel in the 3D images. After the algorithm has converged, the resulting registration parameters represent the initial estimate for the next level, where every 27^{th} voxel is processed, then every 9^{th} voxel, every 3^{rd} and finally every voxel in the MRI image volume.

Multigrid matching is usually motivated by the significant computational gain obtained in 3D registrations. As noticed by several authors,¹³ 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 towards 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. Grey-level normalization of the two 3D images.
3. Pre-registration 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 non-parametric unsupervised algorithm¹⁴ to eliminate the noisy background, followed by a simple region growing algorithm.

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

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

where \bar{I}_{ref} and \bar{I}_{reg} denote the expected values of images $I_{ref}(\cdot)$ and $I_{reg}(\cdot)$, respectively. The number of grey levels of the two volumes is also normalized to the same number G of grey levels, typically $G = 256$.

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 in 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 or multimodal images (the cost functions used in our experiments are presented in the next section):

$$\Theta^* = \underset{\Theta}{\operatorname{argmin}} [E(T_\Theta(x))] \quad , \quad (1)$$

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 axis and the Euler rotation angles $(\hat{\theta}_x, \hat{\theta}_y, \hat{\theta}_z)$.

Fast stochastic optimization as well as a deterministic, gradient-descent algorithms, are used to perform this optimization step, as described in section 2.3.

2.1. 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*.^{4,5} This similarity measure assumes that the two registered images differ only by an additive Gaussian noise,² leading to the following cost function:

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

- In the multimodal case, the inter-image uniformity similarity measure⁸ has gained increasing popularity in the 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 grey 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 , $g = 1, \dots, G$ and the standard deviations σ_g , $g = 1, \dots, G$ of each segmented region are computed. If the two images are correctly registered, the uniformity hypothesis implies that the ratio $\frac{\sigma_g}{\mu_g}$ is minimum over the entire volume.⁸ The following *image uniformity cost function* is thus obtained:

$$E(T_{\Theta}(x)) = \sum_{g=1}^G \frac{N_g \sigma_g(T_{\Theta}(x))}{N \mu_g(T_{\Theta}(x))} , \quad (3)$$

where:

$$\mu_g(T_{\Theta}(x)) = \frac{1}{N_g} \sum_{x|I_{ref}(x)=g} I_{reg}(T_{\Theta}(x)) , \quad (4)$$

and:

$$\sigma_g(T_{\Theta}(x)) = \sqrt{\sum_{x|I_{ref}(x)=g} [I_{reg}(T_{\Theta}(x)) - \mu_g(T_{\Theta}(x))]^2} . \quad (5)$$

In (3), 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.

- Finally, registration by maximization of the mutual information has been proposed recently.⁹ The technique is also based on the same partitioning as in equation 3. The assumption is that the the mutual information I is maximum if the two images are correctly registered :

$$I = \sum_{g=1}^G \sum_{k=1}^K p_{GK}(g, k) \log \frac{p_{GK}(g, k)}{p_G(g)p_K(k)} \quad (6)$$

where G and K stand for the number of grey levels of I_{ref} and I_{reg} . The joint probabilities $p(g, k)$ are the elements of the cooccurrence matrix of $I_{ref}(x)$ and $I_{reg}(T_{\Theta}(x))$ and $p(g)$ and $p(k)$ are the marginal probabilities of $I_{ref}(x)$ and $I_{reg}(T_{\Theta}(x))$, both computed from the normalized histograms of the two images.

2.2. Robust Similarity Measures

As already stated, the least squares criterion (2) works at best under additive Gaussian noise assumptions.² It is well known that least squares are sensitive to gross differences in images due to lesion evolution, to incomplete images, to non Gaussian noise or “outliers”. Outliers generally contribute too much to the overall solution since outlying points are assigned a high weight by the quadratic estimator (2). This remark also holds for the image uniformity cost function (3), 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 mis-registrations 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.^{15,16} 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,¹² applied by Alexander *et al.*,¹¹ 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 algorithm¹⁷ has also 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, that have attractive properties (i.e., satisfactory breakdown points and moderate computational cost), have been widely used in computer vision.¹⁶ This class of robust estimators reduces the optimization problem to a simple, weighted least squares problem.¹⁵ They have a theoretical breakdown point of $\frac{1}{p+1}$, where p is the number of parameters to estimate.¹⁵

In our case, a robust M-estimation of parameters Θ is obtained by introducing a robust error norm ρ in the cost function (2) and (3)¹⁵:

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

$$E(T_{\Theta}(x)) = \sum_x \rho \{I_{ref}(x) - I_{reg}(T_{\Theta}(x)), C\} . \quad (7)$$

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

$$E(T_{\Theta}(x)) = \sum_{g=1}^G \frac{N_g}{N} \widetilde{\sigma}_g(T_{\Theta}(x)) \quad (8)$$

where:

$$\widetilde{\sigma}_g(T_{\Theta}(x)) = \sqrt{\sum_{x|I_{ref}(x)=g} \rho \{I_{reg}(T_{\Theta}(x)) - \widetilde{\mu}_g(T_{\Theta}(x)), C\}} , \quad (9)$$

and

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

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

Let us notice that the non robust cost functions (2) and (3) 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 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 (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 have to be computed using the robust M-estimator.

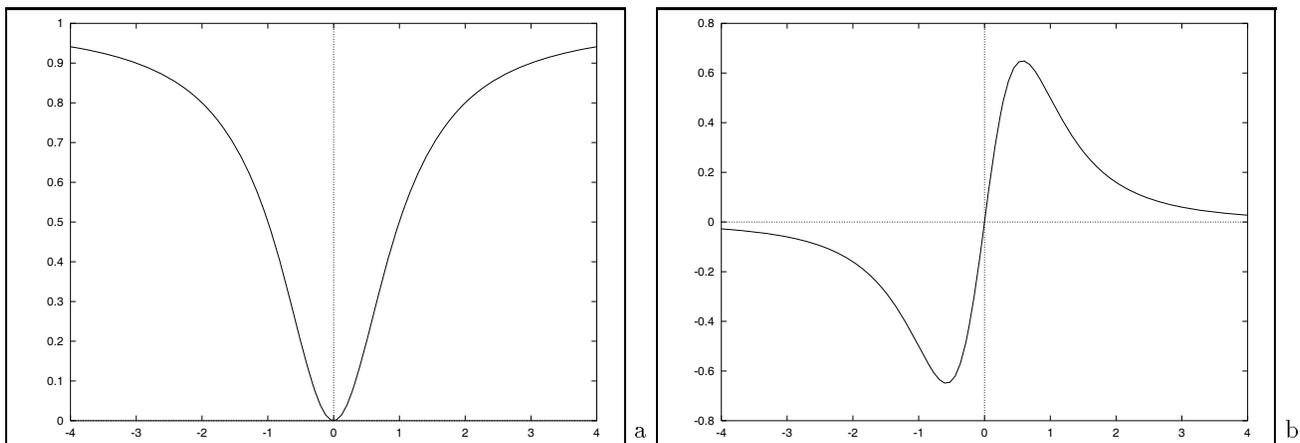
For the experiments presented in this paper we take ρ to be the Geman-McClure estimator (fig. 1):

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

that has successfully been applied in optical flow estimation and for image restoration.¹⁶ Function ρ has a shape that rejects large residual errors. The influence function (fig. 1(b)) is the derivative of function $\rho(x)$ and characterizes the influence of the residuals. As can be seen on 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:

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

Figure 1. The Geman-McClure robust estimator $\rho(x)$ (a) and its influence function (b) for $C = 1$



can be viewed as outliers, as the outliers rejection begins where $\frac{\partial^2 \rho}{\partial x^2} = 0$. We have also experimented with the Tukey “biweight” estimator as well as with the truncated least squares robust function.¹⁶ We privileged the implementation of the Geman-McClure estimator because it requires less calculations for almost the same accuracy with regard to the Tukey “biweight” and is more efficient than the truncated least squares.

The calculation of the registration parameters Θ involves the minimization of the non-linear 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 $C = \alpha.C$ 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 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 during the initial segmentation procedure. The two strategies provides us with approximately the same results.

2.3. Fast Stochastic Optimization

The robust and non robust cost functions considered previously all lead to highly non linear estimation problems, involving many local minima. In most image registration methods involving the minimization of a cost function, deterministic optimization algorithms are applied. They are known to be very sensitive to local minima of the cost function, unless they are initialized close to the optimal solution.

The fast stochastic optimization process, which is applied here is far less sensitive to local minima, yielding better, often close to the optimum, solutions. The optimization technique used in our implementation relies on an iterative fast simulated annealing algorithm¹⁸ based on the Gibbs sampler dynamics.¹⁹ 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.¹⁸

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).²⁰ 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, a multigrid data processing is implemented, as explained previously.

2.4. 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 revealed to be fast and efficient.

2.5. 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 Tesla Bruker system. Each MR image set was obtained with a multi-slice multi-echo sequence (echo time TE=15ms, repetition time TR=2000ms). The size of the images was 256×256 (FOV=25.6cm) with pixel size of $1\text{mm} \times 1\text{mm}$ and a slice thickness of 5mm.
- 3D MR scans were acquired on partially epileptic patients using a 2 Tesla Bruker system. The images were obtained with a gradient echo sequence (TR=40ms and TE=8.5ms, flip angle=45 deg), image size is $128 \times 128 \times 128$ (FOV=25.6cm).
- SPECT imaging was performed on a double-headed camera (Elsint Helix) with low-energy and high resolution parallel-hole collimators, using 700MBq of $^{99m}\text{Tc} - \text{HMPAO}$ or $^{99m}\text{Tc} - \text{ECD}$. The camera was operated in the “stop and shoot” mode with acquisition at 3 deg intervals, acquiring 120 views at 40s per interval (120 projections, 64×64 matrix). Slices were reconstructed with a 128×128 matrix. System resolution was measured at 8mm full width at half maximum (FWHM) in all planes at the center of the field.

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

To test the registration algorithms with ground truth data, a 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. Thanks to 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 a 2D-3D image analysis software (MEDIMAX) developed at the IPB. This software, running under Unix, is 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>).

3. RESULTS

We have compared the robust least-squares (RLS) (7) and robust image uniformity (RIU) (8) approaches to the standard least-squares (LS) method (2) and to the image uniformity (IU) technique (3) and to the recently proposed mutual information criterion (MI).⁹

3.1. Single Modal Registration

Registration experiments were performed both with 2D and 3D images. A first class of experiments consisted in applying a known transformation (translations and rotations) to a set of MRI slices or volumes to create a second image set. 25% of the transformed images was then corrupted by salt and pepper noise, to simulate outliers. For each method, the estimated registration parameters were compared to the true ones 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 $+30$ degrees. As we can see in Table 1, the robust algorithms achieved subvoxel registration errors while the non robust (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 2(c) shows an example where the standard method (LS) failed to correctly register the MR slices shown in figures 2(a) and 2(b), but where the RLS achieved accurate matching by discarding the outliers. The registration error shown in figures 2(c) and 2(d) is the squared image difference after registration. The registration errors 2(c) and 2(d) are normalized to the maximum display value for better visualization.

Figure 2. 2D robust registration. (a) Reference image. (b) Image in (a) rotated by 20 deg, translated by 10 pixels along the x-axis, 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 non robust technique is applied. (d) Difference between the noise free registered image and the image in (a) using the Geman-McClure robust estimation function.

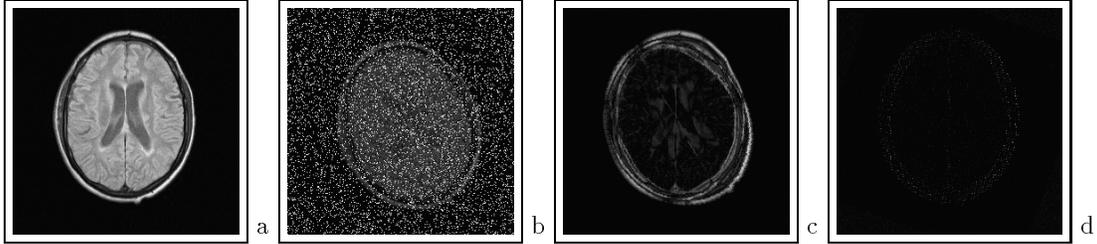


Table 1. 3D registration results. A set of 3D image volumes was artificially transformed using 20 different rigid transformations and the images were corrupted at 25% by salt and pepper noise. The average and the standard deviation of the registration errors computed from the 20 registrations are presented for the different approaches. The translation error is given in voxels and the rotation error in degrees.

Single Modal Registration (3D)						
Approach	Δt_x	Δt_y	Δt_z	$\Delta \theta_x$	$\Delta \theta_y$	$\Delta \theta_z$
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

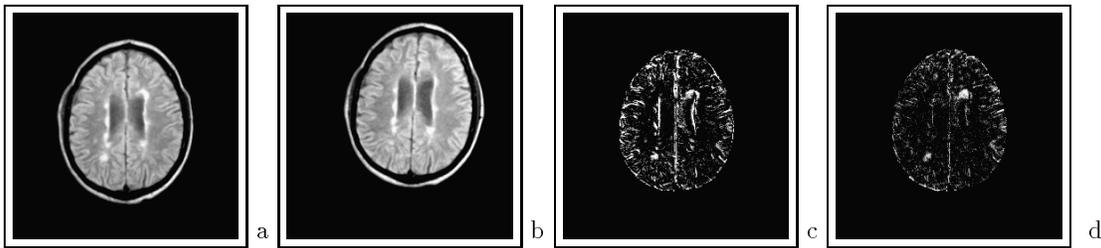
Complementary experiments, with known ground truth, were obtained with a 2D test object, acquired under 35 different rigid transformations by modifying the read and phase gradients during acquisition, as explained previously. Table 2 presents the registration errors for the different techniques in this 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 Tables 1 and 2 (as already noticed, they have rather been devised for multimodal images).

Finally, we have applied the RLS algorithm to a set of MRI slices of a multiple sclerosis (MS) patient, acquired at different dates. 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. 3(c)), are now identified by simple image subtraction (see fig. 3(d)). The robust algorithm achieved better registration than the standard one (fig. 3). Fewer registration artifacts are observed on the cortical sulcus, the falx and the periventricular hyperintensities and multiple sclerosis lesions evolution is clearly displayed. The robust registration technique allowed better follow-up of the disease.

Table 2. 2D Test object registration results. 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.

Approach	Δt_x	Δt_y	$\Delta \theta$
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

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



3.2. Multimodal Registration

To evaluate the multimodal image registration algorithms, a 3D SPECT image volume has been 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 method⁹ which is also suited to multimodal image registration. Table 3 illustrates the robustness of our technique to outliers. The error for the RIU method is around 1 pixel for the translation and 1 degree for the rotation. This is significantly more accurate than 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.

Finally, figure 4 shows a real example of a patient's SPECT image volume (interictal) registered with respect to its MRI counterpart by the robust algorithm. After robust registration of the ictal SPECT volume to the same MRI, the SPECT hyperintensity region (difference between ictal and interictal) has been segmented and superimposed onto the MR image (fig. 5).

Table 3. 3D MRI/SPECT registration results. A set of 3D SPECT image volumes manually pre-registered 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 voxels and the rotation error in degrees.

Approach	Δt_x	Δt_y	Δt_z	$\Delta \theta_x$	$\Delta \theta_y$	$\Delta \theta_z$
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

Figure 4. 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.

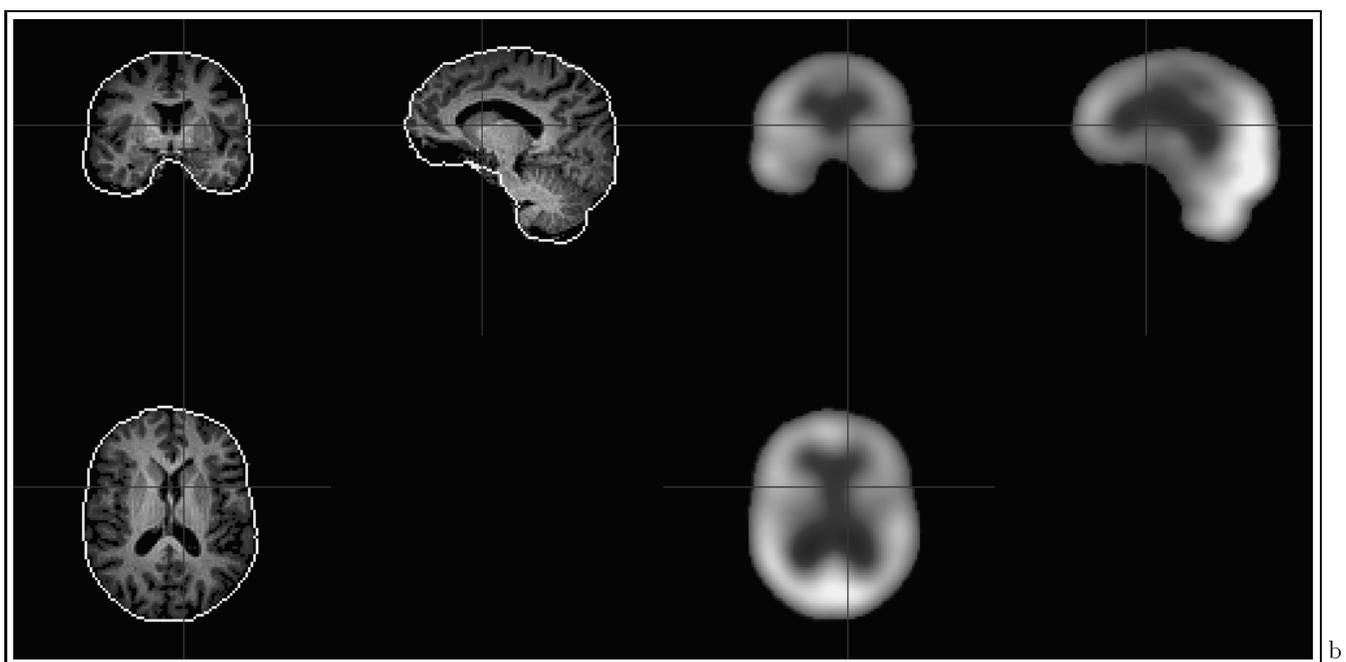
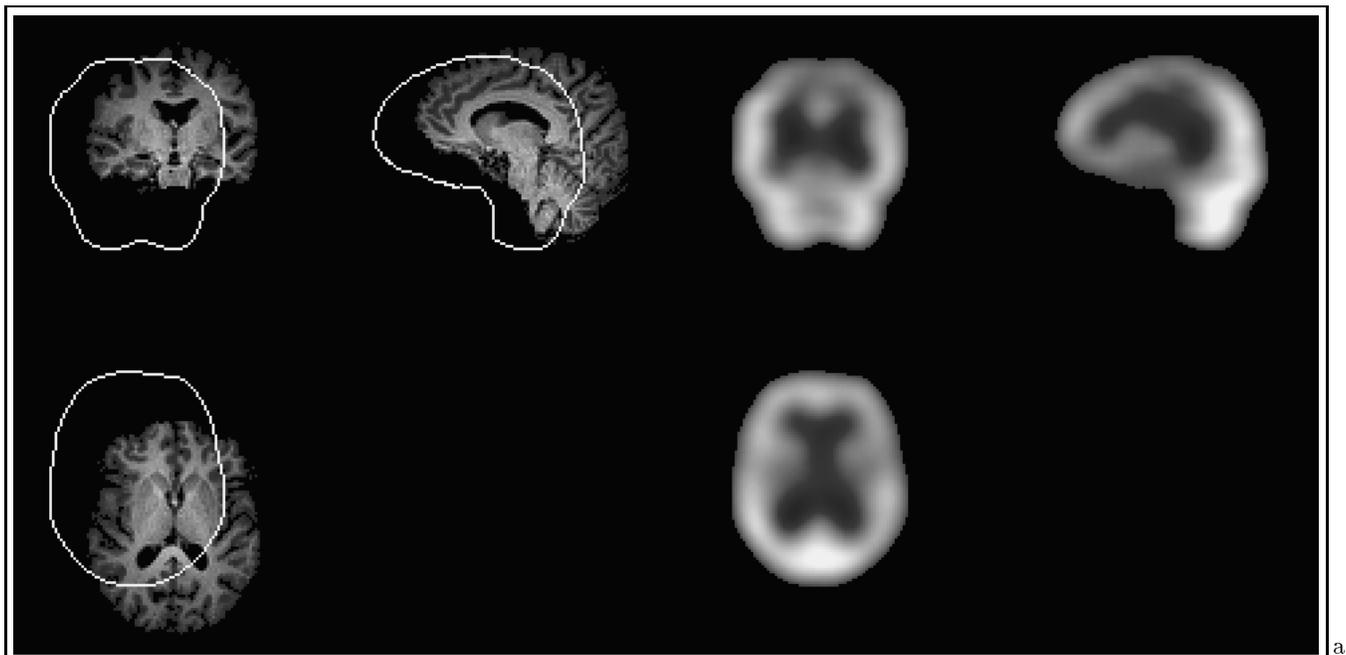
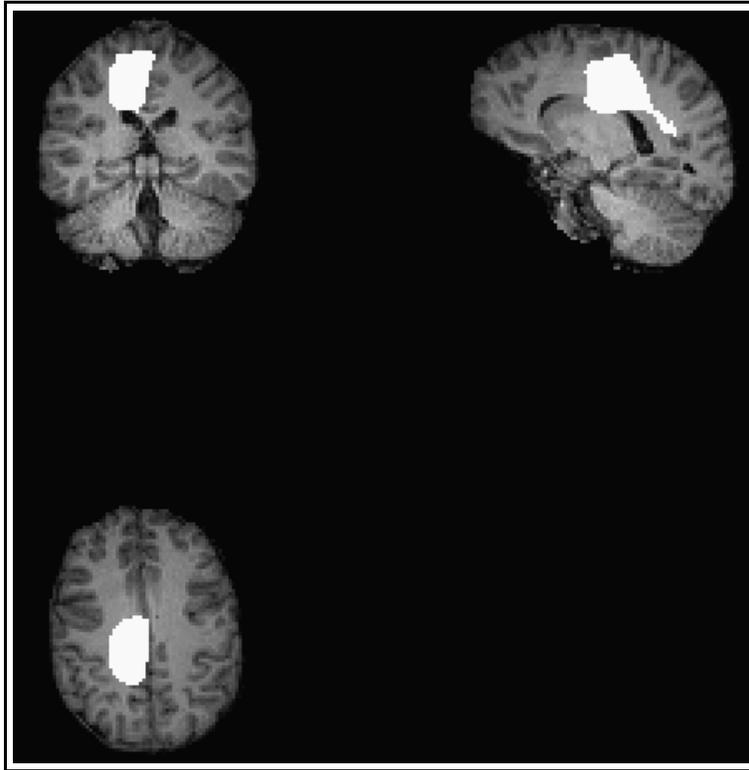


Figure 5. 3D MRI/SPECT representation of a patient presenting partial complex seizures of right temporal origin: the difference image (ROI) between ictal SPECT and interictal SPECT, demonstrating areas of increased perfusion, is shown superimposed onto the corresponding MR image



4. DISCUSSION AND CONCLUSION

The registration methods described in this paper were motivated by the algorithm proposed by^{3,4} in the case of single modality medical image registration and by the model proposed in⁸ in the case of multimodal image matching. These approaches have been improved by the non straightforward 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, due to lesion evolution etc. are taken into account efficiently by robust estimation techniques. The robust functions decide whether a measure is an outlier or not. 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 technique⁹ and to the commonly used image uniformity algorithm.^{3,8} 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. 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 40 min and the RIU method needs 1h 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.

ACKNOWLEDGEMENTS

The study has been supported by the Commission of the European Communities, DG XII, in the framework of the TMR program (Training and Mobility of Researchers), contract Nr ERBFMIBCT960701 and by the “Groupement d’Intérêt Scientifique” (CNRS, CEA, INRIA, MENESR) “Sciences de la Cognition”.

REFERENCES

1. L. G. Brown, “A survey of image registration techniques,” *ACM Computing Surveys* **24**(4), pp. 325–376, 1992.
2. G. Christensen, M. Miller, M. Vannier, and U. Grenander, “Individualizing neuro-anatomical atlases using a massively parallel computer,” *IEEE Computer*, pp. 32–38, January 1996.
3. R. P. Woods, S. R. Cherry, and J. C. Mazziota, “Rapid automated algorithm for aligning and reslicing PET images,” *Journal of Computer Assisted Tomography* **16**(4), pp. 620–633, 1992.
4. J. Hajnal, N. Saeed, E. J. Soar, A. Oatridge, I. R. Young, and G. M. Blyder, “A registration and interpolation procedure for subvoxel matching of serially acquired MR images,” *Journal of Computer Assisted Tomography* **19**(2), pp. 289–296, 1995.
5. N. M. Alpert, D. Berdichevsky, Z. Levin, E. D. Morris, and A. J. Fischman, “Improved methods for image registration,” *Neuroimage* **3**, pp. 10–17, 1996.
6. P. Van den Elsen, J. B. A. Maintz, E. J. D. Pol, and M. A. Viergever, “Automatic registration of CT and MR brain images using correlation of geometrical features,” *IEEE Transactions on Medical Imaging* **14**(2), pp. 384–396, 1995.
7. C. A. Pelizzari, G. T. Y. Chen, D. R. Spelbring, R. R. Weichelbaum, and C. T. Chen, “Accurate three-dimensional registration of CT, PET and/or MR images of the brain,” *Journal of Computer Assisted Tomography* **13**(1), pp. 20–26, 1989.
8. R. P. Woods, J. C. Mazziota, and S. R. Cherry, “MRI-PET registration with automated algorithm,” *Journal of Computer Assisted Tomography* **17**(4), pp. 536–546, 1993.
9. W. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, “Multimodal volume registration by maximization of mutual information,” *Medical Image Analysis* **1**(1), pp. 33–51, 1996.
10. M. Herbin, A. Venot, J. Y. Devaux, E. Walter, F. Lebruchec, L. Dubertet, and J. C. Roucaïrol, “Automated registration of dissimilar images: application to medical imagery,” *Computer Vision, Graphics and Image Processing* **47**, pp. 77–88, 1989.
11. M. Alexander and R. L. Somorjai, “The registration of MR images using multiscale robust methods,” *Magnetic Resonance Imaging* **14**(5), pp. 453–468, 1996.
12. P. J. Rousseeuw, “Least median of squares regression,” *Journal of the American Statistical Association* **79**, pp. 871–880, 1984.
13. F. Heitz, P. Perez, and P. Bouthemy, “Multiscale minimization of global energy functions in some visual recovery problems,” *Computer Vision, Graphics and Image Processing : Image Understanding* **59**(1), pp. 125–134, 1994.
14. N. Otsu, “A threshold selection method from grey-level histograms,” *IEEE Transactions on Systems, Man and Cybernetics* **9**(1), pp. 62–66, 1979.
15. P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim, “Robust regression methods for computer vision: a review,” *International Journal of Computer Vision* **6**(1), pp. 59–70, 1990.
16. M. J. Black and A. Rangarajan, “On the unification of line processes, outliers rejection and robust statistics in early vision,” *International Journal of Computer Vision* **19**(1), pp. 57–91, 1996.
17. C. V. Stewart, “MINPRAN: a new robust estimator for computer vision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(10), pp. 925–938, 1995.
18. E. H. L. Aarts and P. J. M. V. Laarhoven, “Statistical cooling: A general approach to combinatorial optimization problems,” *Philips Journal of Research* **40**(4), pp. 193–226, 1985.
19. S. Geman and D. Geman, “Stochastic relaxation, Gibbs distribution and the bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(6), pp. 721–741, 1984.
20. J. Besag, “On the statistical analysis of dirty pictures,” *Journal of the Royal Statistical Society* **48**(3), pp. 259–302, 1986.