Robust Registration of Dissimilar Single and Multimodal Images

Christophoros Nikou[1,2], Fabrice Heitz[1], and Jean-Paul Armspach[2]

[1] LSIIT UPRES-A CNRS 7005, Universitè Strasbourg I
4 boulevard S. Brant, 67400 Illkirch, France
nikou@mondrian.u-strasbg.fr, Fabrice.Heitz@ensps.u-strasbg.fr
[2] Institut de Physique Biologique UPRES-A CNRS 7004
Facultè de Mèdecine, Universitè Strasbourg I
4, rue Kirschleger, 67085 Strasbourg, France
armspach@alsace.u-strasbg.fr

Abstract. In this paper, we develop data driven registration algorithms, relying on robust pixel similarity metrics, that enable an accurate (subpixel) rigid registration of dissimilar single and multimodal 2D/3D images. A "soft redescending" estimator is associated to a top down stochatic multigrid relaxation algorithm in order to obtain robust, data driven multimodal image registrations. With 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 are compared to the most popular standard similarity metrics, on synthetic as well as on real world image pairs showing gross dissimilarities. Two case-studies are considered: the registration of single and multimodal 3D medical images and the matching of multispectral remotely sensed images showing large overcast areas.

1 Introduction

Although a large variety of image registration methods have been proposed in the literature, only a few techniques have attempted to address the registration of images showing gross dissimilarities. If the case of single modal dissimilar images has been considered in [1], to our knowledge, no specific model has been proposed to handle *multimodal images* exhibiting large dissimilarities. The problem is indeed particularly difficult for multimodal images, showing both localized changes that have to be detected [2] and an "overall" difference (due to differences in the characteristics of the scene observed by multiple sensors). Medical imaging, with its wide variety of sensors (thermal, ultrasonic, X-Ray, MRI and nuclear) is probably one of the first application field, as are remote sensing, military imaging (visible, IR or radar) and multisensor computer vision.

In the present paper, we develop data driven registration methods, relying on pixel (or voxel) similarity metrics, that enable an accurate (subpixel) rigid registration of dissimilar single or multimodal 2D/3D images. Gross dissimilarities are handled by considering similarity measures related to robust M-estimators. In

H. Burkhardt, B. Neumann (Eds.): Computer Vision – ECCV '98,Vol II, LNCS 1407, pp. 51–66, 1998.

particular, a novel pixel similarity metric is proposed for the multimodal case. This metric has shown very efficient for the registration of highly dissimilar images, on which conventional techniques fail. An example of such a multimodal image pair is given Fig. 5, showing two satellite images of France, taken at different optical wavelengths and at different dates. Gross dissimilarities, due to the presence of large overcast areas may be observed (Fig. 5(b)). Subpixel registrations have been obtained in this case (see Section 4).

The remainder of this paper is organized as following. Background and related approaches are presented in Section 2. In Section 3, we introduce two robust similarity metrics for the registration of single and multimodal images. The data-driven registration algorithm, based on these robust similarity measures, is described in the same section. In Section 4, the robust similarity metrics are compared to the most popular standard similarity metrics, on synthetic as well as on real world image pairs showing gross dissimilarities. The registration accuracy is evaluated for two case-studies: the registration of single modal (MRI/MRI) and multimodal (MRI/SPECT) 3D medical images and the matching of multispectral (visible/IR) satellite images showing large overcast areas. The proposed robust similarity measures compare favourably with all standard (non robust) techniques (including the quadratic similarity measure and the multimodality registration criterion devised by Woods *et al.* [3]). The multimodal robust similarity metrics shows also (excepted for one particular case) better performances than the recently proposed mutual information criterion [4,5], that has been recognized as the most efficient method in several recent studies.

2 Background and Standard Similarity Measures

A complete review of standard registration techniques may be found in [6], a classification in [7] and a comparison in [8]. Similarity measure-based approaches rely on the minimization of cost functions that express the pixel or voxel similarity of the images to be aligned. They have been proposed for both single and multimodal image registration [4,5,9,10,11,12]. Similarity metrics for the registration of 2D single modal images, that are to a certain extent robust to image changes have been described by Herbin *et al.* in [1]. Herbin *et al.* make use of deterministic and stochastic sign change criteria to obtain robust registrations of medical image sequences in critical situations corresponding for instance to lesion evolutions [1]. Contrary to the metrics described below, this method does not handle the case of multimodal images.

In this section we briefly present the most popular similarity metrics and describe their limitations. These similarity metrics will be compared, in Section 4, to the robust metrics we propose.

Pixel (or voxel) similarity metric-based registration consists in estimating the parameters Θ of the rigid transformation T_Θ minimizing a cost function $E\left(I_{ref}(.), I_{reg}(T_\Theta(.))\right)$, that expresses the similarity between the single or multimodal image pair:

$$\Theta^* = \arg\min_{\Theta} \left[E\left(I_{ref}(.), I_{reg}(T_\Theta(.))\right)\right] , \tag{1}$$

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 $(\hat{\theta}_X, \hat{\theta}_Y, \hat{\theta}_Z)$, $I_{ref}(.)$ represents the reference image and $I_{reg}(.)$ the image to be registered.

The classical quadratic similarity metric assumes that the two registered images differ only by an additive Gaussian noise [11], leading to the following *least squares cost function*:

$$E\left(I_{ref}(.),\, I_{reg}(T_\Theta(.))\right) \;=\; \sum_x [I_{ref}(x) - I_{reg}(T_\Theta(x))]^2 \;. \tag{2}$$

where x designates the pixel (or voxel) coordinates. Quadratic similarity metrics are related to gaussian sensor models [11], which do not take into account the interimage dissimilarities that may occur in real world applications.

A popular similarity measure for the registration of multimodal image pairs (widely used in medical imaging) is the multimodality similarity metric devised by Woods *et al.* [3]. The fundamental assumption related to Woods criterion is that a uniform region in the reference image corresponds, after registration, to a region that is also uniform in the second image (inter-image uniformity hypothesis).

The reference image is thus first partitioned into G grey level classes, where G denotes the number of grey levels of this image. The resulting spatial partition is projected on the image to be registered, yielding the same partition of this second image. The expected values $\mu_g,\ g = 1, ..., G$ and the standard deviations $\sigma_g,\ g = 1, ..., G$ of each segmented region in the second image are then computed. If the two images are correctly registered, Woods assumes that the normalized variance $\frac{\sigma_g}{\mu_g}$ is minimum over the entire image [3]. The following *inter-image uniformity cost function* is thus defined:

$$E\left(I_{ref}(.),\, I_{reg}(T_\Theta(.))\right) \;=\; \sum_{g=1}^{G} \frac{N_g}{N} \frac{\sigma_g(T_\Theta(.))}{\mu_g(T_\Theta(.))} \;, \tag{3}$$

where:

$$\sigma_g(T_\Theta(.)) = \sqrt{\sum_{x|I_{ref}(x)=g} [I_{reg}(T_\Theta(x)) - \mu_g(T_\Theta(.))]^2} \;, \tag{4}$$

and:

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

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.

As pointed out by Woods [3], the inter-image uniformity hypothesis may only be a crude approximation when gross dissimilarities are present in the multimodal image pair. This is always the case when the multimodal pair is used for the complementary and non redundant information one image provides with respect to the other.

We finally consider the criterion based on the maximization of the mutual information proposed recently and independently in [4,5]. This criterion is based on the same partitioning as in equation (3). The assumption is that the the mutual information is maximum when the two images are correctly registered, yielding the following *mutual information cost function* [4,5]:

$$E\left(I_{ref}(.),\, I_{reg}(T_\Theta(.))\right) \;=\; -\sum_{g=1}^{G}\sum_{k=1}^{K} p(g,k)\log\frac{p(g,k)}{p(g)p(k)} \tag{6}$$

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 $I_{ref}(.)$ and $I_{reg}(T_\Theta(.))$ and $p(g)$ and $p(k)$ are the marginal probabilities of $I_{ref}(.)$ and $I_{reg}(T_\Theta(.))$, both computed from the normalized histograms of the two images.

This criterion has been recognized, in several recent studies, as yielding the best results in multimodal medical image registration. It will be compared to our robust multimodal registration criterion in Section 4.

3 Robust Similarity Metrics-Based Registration

3.1 Robust similarity measures

Standard similarity-based approaches do not model the information differences between images in a single or multimodal pair and, as a consequence, are not robust with respect to them. To increase robustness, the cost function must thus be forgiving about outlying measurements.

Robust estimators have become popular in computer vision applications because they have proven effective in tolerating gross outliers in data [13,14]. A review on robust estimators in computer vision may be found in [13]. A collection of non linear robust estimators, including least median of squares, least trimmed squares, M-estimators, Hough transforms, RANSAC and MINPRAN algorithms are presented in [15,14]. The robustness of these estimators to situations in which mixture of data from multiple (coherent) structures plus gross outliers are to be handled is studied in depth in [14]. Stewart [14] shows that the estimated parameters may be heavily skewed in such situations.

In the following we consider the class of M-estimators [16] that has shown attractive properties (i.e., satisfactory breakdown points and moderate computational cost) in computer vision applications [15,17]. This class of robust estimators reduces the optimization problem to a simple, low cost, weighted least squares problem, as explained in [15,13]. A robust M-estimator of parameters Θ is obtained by introducing a robust error norm ("loss" function) ρ in the similarity metrics (2) and (3) [13].

For the single modality case, we consider the now standard *robust least squares cost function*:

$$E\left(I_{ref}(.),\, I_{reg}(T_\Theta(.))\right) \;=\; \sum_{x}\rho\left\{I_{ref}(x) - I_{reg}(T_\Theta(x)), C\right\}\;. \tag{7}$$

where C is a scale (noise) parameter and ρ is a non quadratic error norm (penalty function) associated with the M-estimator. Variants of this robust cost function have been used with success in image processing and computer vision problems such as surface reconstruction [15], image segmentation, computed imaging, optical flow measurement [18,17], etc.

For multimodal images, we define a *robust inter-image uniformity cost function*:

$$E\left(I_{ref}(.),\, I_{reg}(T_\Theta(.))\right) \;=\; \sum_{g=1}^{G} \frac{N_g}{N} \widetilde{\sigma_g}(T_\Theta(.)) \tag{8}$$

where:

$$\widetilde{\sigma_g}(T_\Theta(.)) \;=\; \sqrt{\sum_{x|I_{ref}(x)=g} \rho\left\{I_{reg}(T_\Theta(x)) - \widetilde{\mu_g}(T_\Theta(.)), C\right\}}\;, \tag{9}$$

and

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

Let us notice that the non robust cost functions (2) and (3) correspond to the special case $\rho(x, C) = x^2$ (for defining (8) we consider a non normalized version of (3), which has shown more efficient than the original Woods' criterion). In the single modal case (7), the cost function is simply defined as a 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 take into account gross 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.

For the experiments presented in this paper we have tested two "hard redescending" M-estimators [14] (namely the truncated quadratic ρ-function [15] and the Tukey "biweight" ρ-function), as well as a "soft redescending" estimator (the Geman-McClure ρ-function [15]). We privileged the Geman-McClure estimator because it required less calculations for almost the same accuracy as the Tukey "biweight" estimator. It showed less sensitive to initialization than the truncated quadratic. The Geman-McClure ρ-function [15] is defined by:

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

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.

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. A good strategy [14] consists in starting 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 also estimated C as the noise variance computed on homogeneous regions of the original images (other statistical methods for estimating C from the data may be found in [14]). These different strategies provided us with almost the same qualitative results.

3.2 The Multiresolution Stochastic Registration Algorithm

The robust estimators and the registration criteria considered previously are highly non linear, involving non convex cost functions having multiple local minima [19]. In most image registration methods based on the minimization of a cost function, deterministic optimization algorithms are applied. They are known to be very sensitive to local minima, unless they are initialized close to the optimal solution.

In order to increase robustness to local minima of the similarity function and to obtain data driven registrations, the parameter space has been discretized and a fast stochastic optimization algorithm has been applied. Stochastic optimization, based on random sampling, is far less sensitive to local minima, yielding better, often close to the optimal solutions [14]. The optimization technique used in our implementation is based on the Gibbs sampler [20]. A high value is adopted for the initial temperature in a simulated annealing procedure and a fast exponentially decreasing temperature schedule is considered instead of the optimal logarithmic descent [20]. The solution obtained after a given number of steps is further refined by a deterministic extension of the above algorithm, known as Iterated Conditional Modes (ICM). ICM is a deterministic Gauss-Seidel like algorithm, that only accepts configurations decreasing the cost function. It has fast convergence properties and local minima are not a problem, since the first stochastic optimization step provides a good initialization.

The optimization algorithm was applied on a sequence of multiresolution grids, using a standard top-down approach starting from the coarsest resolution level [17,21]. 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} (16^{th}) voxel (pixel) in the 3D (2D) images. After the algorithm has converged, the resulting registration parameters represent the initial estimate for the next level, where every 27^{th} (8^{th}) voxel is processed, then every 9^{th} (4^{th}), every 3^{rd} (2^{nd}) and finally every voxel (pixel) in the image. The search space and the visited configurations were reduced while the resolution increases in order to gradually fine tune the solutions obtained on the coarser resolution levels. The first grids generally provided a good approximation of the final solution. Multigrid matching is usually motivated by the significant computational gain obtained in the registration. As noticed by several authors

[19], multigrid algorithms are also far less sensitive to local minima in the cost function than single resolution optimization schemes. It has indeed been conjectured that multigrid analysis may, to a certain extent, smooth the "landscape" of the objective function to minimize. This yields fast convergence towards good solutions [19].

4 Experimental Results

Registration experiments were performed with both 2D and 3D images. The following similarity measures have been implemented and compared:

- the standard least-squares (LS) similarity measure (Eq. 2) ;
- the inter-image uniformity (IU) criterion [3] (Eq. 3) ;
- the mutual information (MI) criterion [4,5] (Eq. 6) ;
- the robust least-squares (RLS) similarity metrics (Eq. 7) ;
- the robust inter-image uniformity (RIU) criterion (Eq. 8).

LS and RLS may only be applied to single modal image registration, whereas the other methods (IU, RIU, MI) have been tested both in single and multimodal registration problems. Two representative case studies have been considered: the registration of single modal (MRI/MRI) and multimodal (MRI/SPECT) 3D medical images showing gross outliers or lesion evolution, and the matching of multispectral (visible/IR) remotely sensed images showing large overcast areas.

4.1 Single Modal Image Registration

Medical Images. A first class of experiments consisted in applying a known rigid transformation (3D translation and rotation) to a set of MRI volumes to create a second image set. 25% of the transformed images was then corrupted by salt and pepper noise, to simulate gross outliers (see Fig. 1(a-b)). 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. Let us notice that large rotations are generally difficult to handle with standard, deterministic approaches (in which initializations close to the desired solution are necessary). This is not the case of the stochastic sampling algorithm used here.

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, the "best" method referenced at the present time, also achieved subvoxel registration but its performances are slightly inferior to the results obtained by the RLS technique.

Figure 1(c) shows an example where the standard method (LS) failed to correctly register the MR slices shown in Figures 1(a) and 1(b), but where the RLS achieved accurate matching by discarding outliers. The difference in

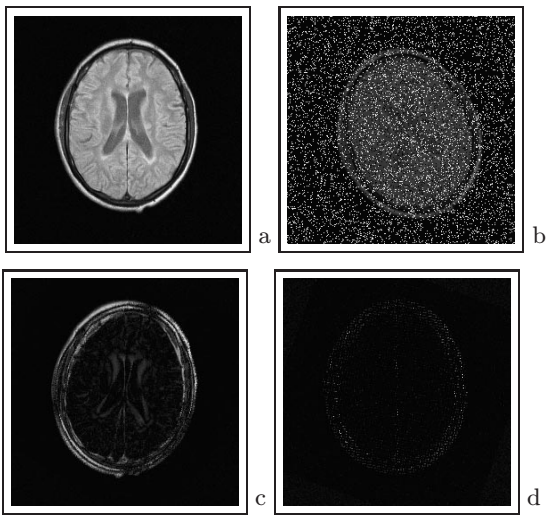

Fig. 1. Robust registration of MR images. **(a)** Reference image. **(b)** Image in (a) rotated by 20 deg, translated by 10 pixels along the x-axis, 10 pixels along the y-axis and corrupted at 25% with salt and pepper noise with large magnitude. **(c)** Difference between the (*noise free*) registered image and the image in (a) (*LS similarity metric*). **(d)** Difference between the (*noise free*) registered image and the image in (a) (*RLS similarity metric*)

accuracy is readily visible on the registration error shown in Figures 1(c) and 1(d).

We also show in Fig. 2 an example of the application of the RLS algorithm to the detection of changes in a set of MRI slices of a multiple sclerosis patient, acquired at different dates. Figure 2 illustrates a case on which small differences due to lesion evolution, which were not well distinguished previously due to misalignment by the standard LS similarity metric (Fig. 2(c)), are now clearly identified by simple image subtraction (Fig. 2(d)). This result has been validated by an expert physician from IPB.

Remotely Sensed Images. Two images of France, in the infra-red band of NOAA (Fig. 3(a-b), acquired at different dates and showing large overcast areas, were manually registered by an expert from LSIIT to establish ground truth. One of the images has been transformed using different 2D rotation and translation parameters and the registration algorithms were applied. This case, contrary to the example considered previously (Section 4.1), does not correspond to a corruption of the data by gross outliers, but to the presence of multiple coherent structures (i.e. ground and clouds) in the data. Mixture of data from multiple (coherent) structures introduces a significant bias in all robust estimators, as shown in a recent study by Stewart [14]. The performances of the robust methods are affected by this bias, as can be seen in Table 2 in which the different approaches

Table 1. Single modal registration of $3D$ MRI images. An MR volume 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. Translation errors are given in voxels and rotation errors in degrees

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

are compared. The registrations are not as accurate as in the previous case, although a subpixel accuracy is reached, and the difference between methods is less pronounced. The robust methods produce nevertheless the best results and compare favourably to the MI approach.

Figure 3 illustrates the contribution of the RLS metric with respect to a non robust LS metric, in the registration of the original infra-red image pair. The original images show a misregistration of about 3 pixels. Clouds in the second image lead the LS technique to a slight misalignment (Fig. 3(c)) while the RLS measure provides a more accurate registration (Fig. 3(d)). The difference is readily visible along the south-west coast of France. The registration errors presented in Figures 3(c-d) are obtained by subtraction of the registered image from the reference image in Figure 3(a), followed by contrast modifications for visualization purpose.

Table 2. Single modal registration of $2D$ remotely sensed infra-red images. Two images of the infra-red electromagnetic band of NOAA satellite acquired at different dates have been manually registered to create ground truth. One of the images has undergone 20 different rigid transformations using different translation and rotation values. The average and the standard deviation of the registration errors are presented for the different approaches. Translation errors are given in pixels and rotation errors in degrees

Approach	Δt_x	Δt_y	$\Delta\hat{\theta}$
LS	0.42 ± 0.18	0.31 ± 0.41	0.32 ± 0.18
IU	0.52 ± 0.21	0.77 ± 0.40	0.30 ± 0.25
MI	0.49 ± 0.54	0.63 ± 0.25	0.75 ± 0.89
RLS	0.36 ± 0.10	0.27 ± 0.37	0.30 ± 0.25
RIU	0.34 ± 0.17	0.70 ± 0.28	0.18 ± 0.13

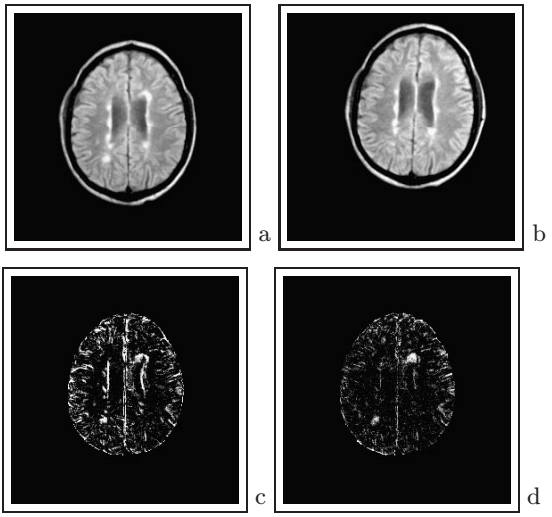

Fig. 2. Change detection in a MRI image sequence. **(a)** Multiple sclerosis patient MR image. **(b)** Image of the same patient acquired several months later. **(c)** Difference between the registered image and the image in (a) (*LS similarity metric*). **(d)** Difference between the registered image and the image in (a) (*robust RLS similarity metric*)

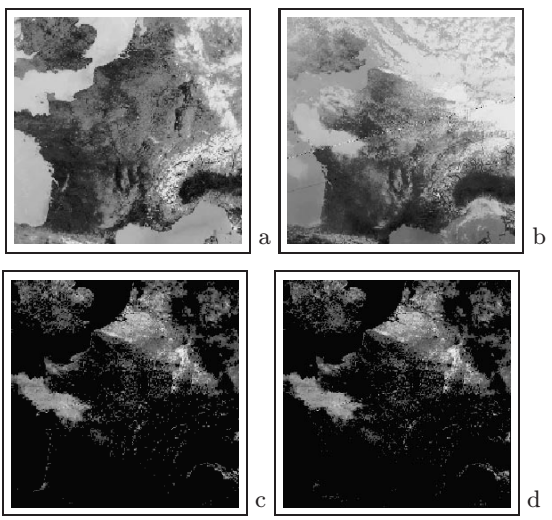

Fig. 3. Single modal registration of remotely-sensed images. **(a)** Image of France in the infra-red band of NOAA (*02/10/97*). **(b)** Image of France in the infra-red band of NOAA (*02/05/97*). **(c)** Registration error (*LS similarity metric*). **(d)** Registration error (*robust RLS similarity metric*)

4.2 Multimodal Registration

Medical Images. To evaluate the ability of the robust similarity metrics to handle multimodal image pairs, a 3D SPECT image volume has been manually registered to its corresponding MRI volume with the aid of an expert physician from IPB. The manually registered SPECT volume was then transformed using the same 3D translation and rotation parameters as in the previously described experiments (Section 4.1). To simulate outliers, 25% of the SPECT image was corrupted by salt and pepper noise. The robust inter-image uniformity technique RIU has been compared to the inter-image uniformity similarity function IU [3] and to the mutual information MI criterion [4,5]. Table 3 shows the robustness of the different similarity measures to gross outliers. The error for the RIU method is about 1 voxel for the translation parameters and 1 degree for the Euler rotation angles. This is significantly more accurate than the IU approach. The proposed robust similarity metric also compares favourably to the MI criterion which yields registrations that are better than the IU criterion but are generally below RIU.

Table 3. Multimodal registration of $3D$ MRI/SPECT images. A $3D$ SPECT image volume 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. Translation errors are given in voxels and rotation errors in degrees

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

Figure 4 shows a real example of a patient SPECT image volume registered with respect to its MRI counterpart by the robust algorithm. The accuracy of the registration has been evaluated by visual inspection and has been considered as satisfactory by an expert.

Remotely Sensed Images. We consider again the case of multispectral remotely sensed images, presenting coherent data corruption due to large overcast areas. Two images, one in the visible and one in the infrared band of NOAA, acquired at different dates (Fig. 5(a-b)) were manually registered to establish ground truth. One of the images has been transformed using different rotation and translation parameters and the multimodality registration algorithms were applied. The performances of the different methods are summarized in Table 4. As expected the robust RIU criterion provides registrations that are significantly more accurate than the non robust IU technique. The difference between

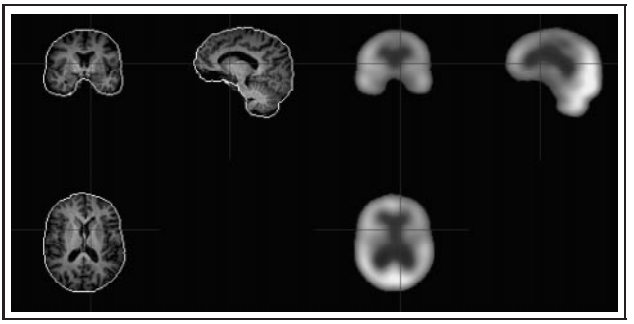

Fig. 4. Robust MRI/SPECT volume registration. The SPECT and MRI volumes with the SPECT contours superimposed on the MRI are shown (*multiplanar visualization*) after robust registration (*RIU similarity metric*)

the tested similarity metrics is however not as pronounced as for the medical images registration problem (in which gross outliers were considered). This may again be explained by the bias introduced by the mixture of data from multiple coherent structures on the robust estimation [14]. In this particular case, the mutual information MI criterion yields, in the average, the best results. Let us however notice that the variance of the MI estimate is significantly higher than the variance of the robust RIU criterion (see Table 4), which tends to temper the conclusion in this case.

Table 4. Multimodal registration of $2D$ visible/infra-red images. Two images, one of the visible and one of the infra-red electromagnetic band of NOAA satellite acquired at different dates have been manually registered to create ground truth. One of the images has undergone 20 different rigid transformations using different translation and rotation values. The average and the standard deviation of the registration errors are presented for the different approaches. Translation error are given in pixels and rotation errors in degrees

Approach	Δt_x	Δt_y	$\Delta\hat{\theta}$
IU	1.34 ± 0.87	1.04 ± 0.34	0.34 ± 0.27
MI	0.40 ± 0.68	0.31 ± 0.74	0.24 ± 0.37
RIU	0.51 ± 0.34	0.76 ± 0.37	0.26 ± 0.20

Figure 5 presents the registration of the original multimodal pair. The images from the NOAA visible band (Fig. 5(a)) and from the NOAA infra-red band (Fig. 5(b)), acquired at different dates have been registered using the IU, RIU and MI approaches. In this particular case, the non robust IU metric and the MI criterion provided the same final registrations. As may be seen Fig. 5 (c), the IU metric, yields a misregistration, that is visible on the error image, along the south-west coast of France. This is not the case of the robust RIU similarity

measure (Fig. 5 (d)) which provides an accurate registration of this dissimilar multimodal image pair. Let us notice that the multimodal registration error shown in Fig. 5 (c-d) is defined as the difference between the registered IR image and the IR image acquired at the same instant as the visible band reference image.

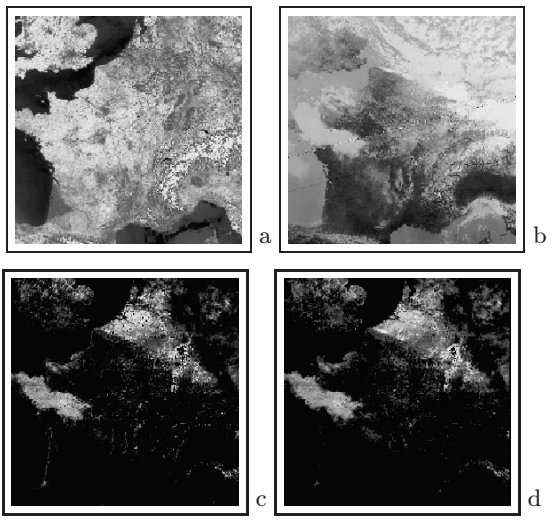

Fig. 5. Multimodal registration of visible/IR remotely-sensed images. **(a)** Image of France in the visible band of NOAA (*02/10/97*) (reference image). **(b)** Image of France in the infrared band of NOAA (*02/05/97*). **(c)** Registration error (*IU similarity metric*). **(d)** Registration error (*robust RIU similarity metric*). The registration error is defined as the difference between the registered IR image and the IR image acquired at the same instant as the visible band reference image (a)

The LS and RLS techniques require approximately the same average computation times: 20 mn cpu time for 3D $128 \times 128 \times 128$ images on a HP 715/80 workstation. For the same data size, the IU method takes 35 mn, the MI technique 40 mn and the RIU method needs 60 mn cpu time. In the case of 2D images (256×256), the RIU metric requires 4 mn cpu time while each of the other techniques takes approximately 1-2 minutes. As can be seen, the additional computational complexity introduced by the robust estimation is acceptable and these methods may thus be used with profit to improve the accuracy in many critical single or multimodal image registration problems.

5 Conclusion

The robust similarity metrics-based registration methods described in this paper were motivated by the lack, in existing approaches, of specific models for gross dissimilarities or outlying data that are often present in single and multimodal image pairs. The proposed stochastic multigrid registration algorithms have two major advantages over standard methods:

- No manual initialization near the optimal solution is required to obtain an accurate registration. Local minima, a major problem in standard image registration techniques, are avoided by the use of fast multigrid random sampling algorithms. This results in a fully data driven method that requires no human interaction.
- Gross image differences are taken into account efficiently by robust M-estimators. To our knowledge, the registration of *multimodal images* showing gross dissimilarities or mixture of data from multiple coherent structures has never been evoked until now.

As a conclusion, let us emphasize that the approach proposed in this paper is comprehensive and not limited to medical or remotely-sensed images. Other potential application fields [6] such as military imaging, multisensor robot vision or the multisource analysis of artistic patrimony [2] may benefit from the robustness of these methods.

Acknowledgments

This 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'Intèrt Scientifique" (CNRS, CEA, INRIA, MENESR) "Sciences de la Cognition".

The authors wish to thank Dr. Denis Bruckert (Groupement de Recherche en Tèlèdètection et Radiomètrie, LSIIT) for providing the NOAA images and Dr. Izzie-Jacques Namer (IPB) for providing and interpreting the MR and SPECT data.

References

1. M. Herbin, A. Venot, J. Y. Devaux, E. Walter, F. Lebruchec, L. Dubertet, and J. C. Roucayrol. Automated registration of dissimilar images: application to medical imagery. *Computer Vision, Graphics and Image Processing*, 47:77–88, 1989. 51, 52, 52
2. F. Heitz, H. Maître, and C. de Couessin. Event detection in multisource imaging: application to fine arts painting analysis. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 38(4):695–704, 1990. 51, 64
3. R. P. Woods, J. C. Mazziota, and S. R. Cherry. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17(4):536–546, 1993. 52, 53, 53, 53, 57, 61

4. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997. 52, 52, 54, 54, 57, 61
5. 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):33–51, 1996. 52, 52, 54, 54, 57, 61
6. L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992. 52, 64
7. P. Van den Elsen, E. J. D. Paul, and M. A. Viergever. Medical image matching - a review with classification. *IEEE Engineering in Medicine and Biology*, 12(1):26–39, 1993. 52
8. J. West, M. Fitzpatrick, M. Wang, B. Dawan, C. Maurer, Jr. R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. Van den Elsen, S. Napel, T. Sumanaweera, B. Harkness, P. Hemler, D. Hill, D. Hawkes, C. Studholme, A. Mainz, M. Viergever, G. Malandain, X. Pennec, M. Noz, G. Maguire, Jr. M. Pollack, C. Pelizzari, R. Robb, D. Hanson, and R. Woods. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21(4):554–566, 1997. 52
9. G. Malandain, S. Fernàndez-Vidal, and J. M. Rocchisani. Improving registration of 3D medical images using a mechanical-based method. *Lecture Notes in Computer Vision (ECCV'94)*, 801:131–136, 1994. 52
10. J. P. Thirion, A. Gourdon, O. Monga, A. Guéziec, and N. Ayache. Fully automatic registration of 3D CAT-scan images using crest lines. 14^{th} *International Conference EMBS'92 (IEEE)*, pages 1888–1890, 1992. 52
11. G.E. Christensen, M.I. Miller, M.W. Vannier, and U. Grenander. Individualizing neuro-anatomical atlases using a massively parallel computer. *IEEE Computer*, pages 32–38, January 1996. 52, 53, 53
12. J. Sato and R. Cipolla. Image registration using multi-scale texture moments. *Image and Vision Computing*, 13(5):341–353, 1995. 52
13. 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):59–70, 1990. 54, 54, 54, 54
14. C. Stewart. Bias in robust estimation caused by discontinuities and multiple structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(8):818–833, 1997. 54, 54, 54, 54, 55, 55, 56, 56, 58, 62
15. 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):57–91, 1996. 54, 54, 54, 55, 55, 55, 55
16. J. Huber. *Robust statistics*. John Wiley and sons, New York, 1981. 54
17. J. M. Odobez and P. Bouthemy. Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4):348–365, 1995. 54, 55, 56
18. M. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, 1996. 55
19. 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):125–134, 1994. 56, 57, 57

20. 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):721–741, 1984. 56, 56
21. D. Terzopoulos. Image analysis using multigrid relaxation methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):129–139, 1986. 56