Journal of Electronic Imaging 8(3), 255–262 (July 1999).

Creaseness-based computer tomography and magnetic resonance registration: Comparison with the mutual information method

David Lloret
Antonio M. López
Joan Serrat
Juan J. Villanueva
Universitat Autònoma de Barcelona
Computer Vision Center and Departament d'Informàtica
Edifici O, 08193-Cerdanyola, Spain
E-mail: david@cvc.uab.es

Abstract. *This paper describes a method which uses the skull as a landmark for automatic registration of computer tomography to magnetic resonance (MR) images. First, the skull is extracted from both images using a new creaseness operator. Then, the resulting creaseness images are used to build a hierarchic structure which permits a robust and fast search. We have justified experimentally the performance of several choices of our algorithm, and we have thoroughly tested its accuracy and robustness against the well-known mutual information method for five different pairs of images. We have found both comparable, and for certain MR images the proposed method achieves better performance.* © *1999 SPIE and IS&T.* [S1017-9909(99)00403-1]

1 Introduction

Image registration or matching attempts to solve the problem that arises when two images taken at different times, by different sensors or from different viewpoints, need to be compared. An upcoming application of image registration is in the field of medical images, especially since the introduction of three-dimensional (3D) modalities. Many methods have been proposed for multisensor medical image registration. In a recent survey, Maintz and Viergever[1] classify an exhaustive list of publications according to nine different criteria. Automatic intrinsic methods can be divided as volume based or as surface based. Volume-based methods statistically compare the similarity between the intensity values of the two images, while surface-based methods match related surfaces previously extracted from the images. West *et al.*[2] evaluated the accuracy of 12 of these methods for six pairs of images [from computer tomography (CT) and positron emission tomography (PET) to three magnetic resonance (MR) settings] from seven patients. As a group, for the data in the study volume-based methods clearly showed better performance, and they have become the *de facto* golden standard.

Paper NMI-02 received Nov. 2, 1998; revised manuscript received Apr. 12, 1999; accepted for publication May 25, 1999.
1017-9909/99/$10.00 © 1999 SPIE and IS&T.

We have developed a surface-based method based on an idea first presented by Petra van den Elsen.[3] We take the skull—a structure visible both in CT and MR brain images—as the reference landmark. The signal produced by the bone is strong in CT, but weak in MR due to the lack of mobile protons. Seen as a landscape, the skull forms a ridge in the CT image, and a valley in the MR image. Moreover, since the skull is an undeformable structure, it is ideal for matching purposes because only rigid transformations need to be considered. Van den Elsen *et al.* and Maintz *et al.* evaluated the performance of several creaseness operators,[4] they compared them to some edgeness operators,[5] and finally they submitted only edgeness operators to be evaluated.

The novelties of our method are the following: (a) our 3D creaseness (ridge and valleyness) operator is able to extract neatly the skull from both images, with almost no spurious response (Sec. 2), (b) we have added surface matching using chamfer distances as in Ref. 6 to improve robustness and (c) we have investigated experimentally the performance of our method for different choices of settings (Sec. 5). We describe our hierarchical optimization method in Sec. 3, and we shortly review the mutual information paradigm in Sec. 4. In Sec. 5 we compare it to the mutual information method. Conclusions, results and future work are discussed in Sec. 6.

2 Creaseness Operator

Today's literature defines a crease (ridge or valley) in many different ways.[7] Among them, the definition based on level set extrinsic curvature (LSEC) is one of the more useful because of its invariance properties. Given a function $L:\mathbf{R}^d \rightarrow \mathbf{R}$, the level set for a constant l consists of the set of points $\{\mathbf{x}|L(\mathbf{x})=l\}$. The simplest situation occurs at $d=2$, when L can be thought of as a topographic relief or landscape and the level sets are its level curves. It is well known that negative minima of the level curve curvature κ, level by level, form valley-like curves and positive maxima

ridge-like curves. In 3D we have level surfaces and the straightforward extension of κ is two times the Mean curvature κ_M of the level surfaces

$$\kappa_M=[2(L_xL_yL_{xy}+L_xL_zL_{xz}+L_yL_zL_{yz})-L_x^2(L_{yy}+L_{zz}) -L_y^2(L_{xx}+L_{zz})-L_z^2(L_{xx}+L_{yy})]2(L_x^2+L_y^2+L_z^2)^{-3/2}. \quad (1)$$

However, the way LSEC is usually computed—directly discretizing Eq. (1) or the equivalent expression in the two-dimensional (2D) case for κ—gives rise to a number of discontinuities: gaps appear at places where we would not expect any reduction of creaseness because they are at the center of elongated objects and creaseness is a measure of medialness for gray-level elongated objects. We have checked that this happens independently of the scheme of discretizing derivatives and, even when the gradient magnitude is far away from the machine's zero, that is, at pixels where Eq. (1) is well defined.

To avoid discontinuities we present an alternative way of computing creaseness from the level set extrinsic curvature. For a d-dimensional vector field $\mathbf{u}:\mathbf{R}^d\rightarrow\mathbf{R}^d, \mathbf{u}(\mathbf{x})=(u^1(\mathbf{x}),\ldots,u^d(\mathbf{x}))^t$, the divergence measures the degree of parallelism of its integral lines

$$\mathrm{div}(\mathbf{u})=\sum_{i=1}^{d}\frac{\partial u^i}{\partial x^i}. \quad (2)$$

Now, if we define the normalized gradient vector field of $L:\mathbf{R}^d\rightarrow\mathbf{R}$ as

$$\overline{\mathbf{w}}=\begin{cases}\mathbf{w}/\|\mathbf{w}\| & \text{if} \quad \|\mathbf{w}\|>0\\ \mathbf{0}_d & \text{if} \quad \|\mathbf{w}\|=0\end{cases} \quad (3)$$

it can be shown that

$$\kappa_d=-\mathrm{div}(\overline{\mathbf{w}}). \quad (4)$$

For $d=3$ Eqs. (1) and (4) is equivalent in the continuous domain. However, and this is the key point, the same approximation using derivatives substituted into Eqs. (1) and (4) gives rise to very different results, namely, Eq. (4) avoids the gaps that Eq. (1) produces on creases. For example, in the case of ridge-like saddle points, as those in Fig. 1(a), κ not only goes down but also suffers a change of sign barrier on the path of the expected ridge-like curve [Fig. 1(c)]. The reason is that the neighborhood of the saddle point is composed mainly of concave zones ($\kappa<0$) but the subpixel ridge-like curve runs through convex zones ($\kappa>0$) missed by the discretization of Eq. (1).

In order to obtain derivatives of a discrete image L in a well-posed manner we have approximated image derivatives by finite centered differences of the Gaussian-smoothed image instead of convolution with Gaussian derivatives of a given standard deviation σ_D. For each pixel, neighbor derivatives are calculated ''on the fly'' as needed and stored in buffers for the next one. This approach is necessary because otherwise the algorithm would require

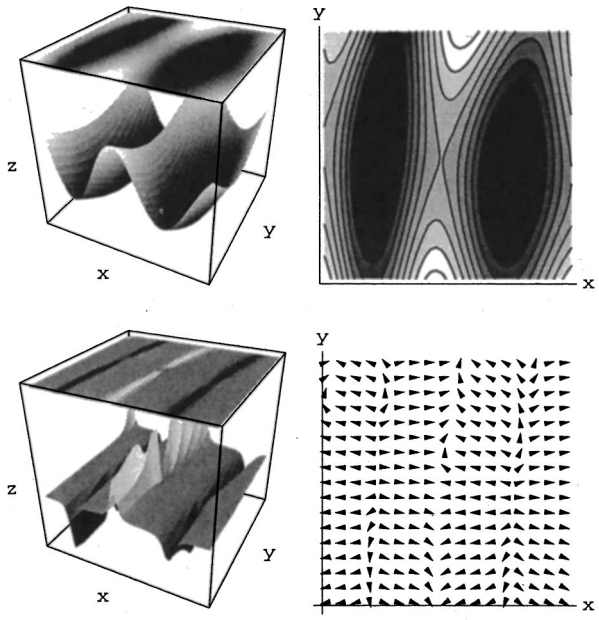


Fig. 1 From top to bottom and left to right: (a) relief with regions having valley-like and ridge-like shape, (b) its level curves, (c) κ (notice the gaps), (d) $\overline{\kappa}$.

ten float images simultaneously. Thus, Eq. (4) produces similar results being computationally less time and memory consuming.

The creaseness measure κ_d can still be improved by pre-filtering the image gradient vector field in order to increase the degree of attraction/repulsion at ridge-like/valley-like creases, which is what κ_d is actually measuring. This can be done by the structure tensor analysis[8,9]

1. Compute the gradient vector field $\mathbf{w}$ and the structure tensor field $\mathbf{M}$

   $$\mathbf{M}(\mathbf{x};\sigma_I)=G(\mathbf{x};\sigma_I)*(\mathbf{w}(\mathbf{x})\cdot\mathbf{w}(\mathbf{x})^t) \quad (5)$$

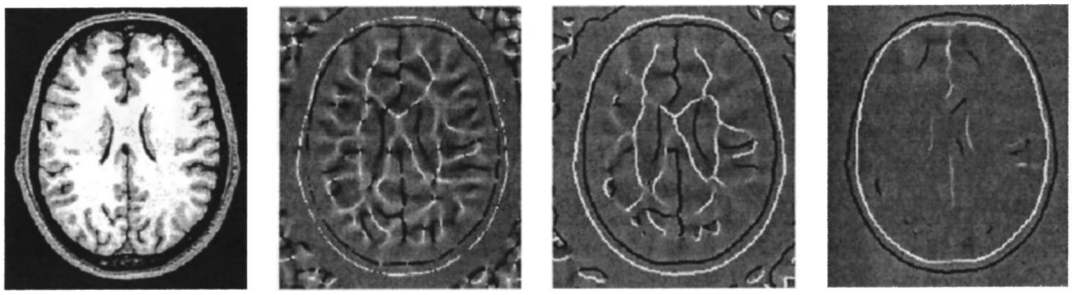
   being $*$ the element-wise convolution of matrix $\mathbf{w}(\mathbf{x})\cdot\mathbf{w}(\mathbf{x})^t$ with the Gaussian window $G(\mathbf{x};\sigma_I)$. The standard deviation σ_I controls the size of the averaging window.

2. Perform the eigenvalue analysis of $\mathbf{M}$. The normalized eigenvector $\mathbf{w}'$ corresponding to the highest eigenvalue gives the predominant gradient orientation. In the structure tensor analysis, opposite directions are equally treated. Thus, to recover the direction we put $\mathbf{w}'$ in the same quadrant in 2D, or octant in 3D, as $\mathbf{w}$. Then, we obtain the new vector field $\tilde{\mathbf{w}}$ and the creaseness measure $\tilde{\kappa}_d$

   $$\tilde{\mathbf{w}}=\mathrm{sign}(\mathbf{w}'^t\mathbf{w})\mathbf{w}', \quad (6)$$

   $$\tilde{\kappa}_d=-\mathrm{div}(\tilde{\mathbf{w}}). \quad (7)$$

   The discretization of $\tilde{\kappa}_d$ results from approximating partial derivatives of Eq. (2) by finite centered differences of the $\tilde{\mathbf{w}}$ vector field. When the smallest neighborhood is taken, namely, the six pixels neighboring a pixel (i,j,k), the following expression is obtained:



(a) (b) (c) (d)

Fig. 2 From left to right: 3D creaseness measures of the MR slice in (a) for $\sigma_D=2.0$ mm: (b) κ, (c) $\tilde{\kappa}$ for $\sigma_I=2.0$ mm, (d) $\mathcal{C}\tilde{\kappa}$ for $c=10\,000$. The gaps in (b) do not appear in (c).

$$\tilde{\kappa}_d[i,j,k]=\tilde{w}^1[i+1,j,k]-\tilde{w}^1[i-1,j,k]+\tilde{w}^2[i,j+1,k]-\tilde{w}^2[i,j-1,k]+\tilde{w}^3[i,j,k+1]-\tilde{w}^3[i,j,k-1]. \quad (8)$$

3. Compute a confidence measure $\mathcal{C}\in[0,1]$ to discard creaseness at isotropic areas. As similarity of the eigenvalues $(\lambda_1,\ldots,\lambda_d)$ of the structure tensor implies isotropy we can base the computation of $\mathcal{C}$ on a normalized measure based on their difference. A suitable choice is

$$\mathcal{C}(\mathbf{x})=1-e^{-(\sum_{i=1}^{d}\sum_{j=i+1}^{d}(\lambda_i(\mathbf{x})-\lambda_j(\mathbf{x}))^2)^2/2c^2}, \quad (9)$$

where threshold c is experimentally chosen. In this way, $\mathcal{C}\tilde{\kappa}_d$ has a lower response than $\tilde{\kappa}_d$ at isotropic regions.

We have tested several methods to compute the eigenvalues. If we only need the highest, the fastest method is Powell's applied to an optimized expression from Ref. 10. If we choose to implement step (3), the fastest is singular value decomposition as implemented in Ref. 11. Figure 2 compares κ, $\tilde{\kappa}$ and $\mathcal{C}\tilde{\kappa}$ on an MR volume.

3 Search Strategy

The former operator is highly suitable for CT-MR registration because it gives a high response in the voxels depicting the skull, which will lead to a quick and accurate alignment of the images. The first step is to reformat the source images into cubic voxels. After the creaseness extraction, the following step is to iteratively transform one of the images until it becomes properly aligned with the other. A suitable function to measure the quality of the alignment is the correlation function $C_T=\Sigma_{\mathbf{x}\in f}f(\mathbf{x})\cdot g(T(\mathbf{x}))$, where f and g are the creaseness images and T represents a transformation whose parameters we want to test. A key step is not to transform all the pixels in the image, but only those with values higher than a small fixed threshold. This step saves up to 95% of the total computations.

The function C_T together with the six parameters of the transformation define a search space which is difficult to optimize because: (a) the function is nonmonotonic, i.e., has many local maxima, (b) the similarity measure is expensive to compute since it involves the transformation of a large 3D image, and (c) translation and rotation parameters cannot be decoupled in order to reduce the dimension of the search space.

An approach to overcome the two first problems is to search within the parameter space at multiple resolutions. As in Ref. 3 we handle multiple resolution by building two pyramids where the CT ridgeness and MR valleyness images are at the bottom and each level is a sampled version of the previous at half resolution, until images have a final size of about 16 pixels in each dimension.

The search starts at the top of the pyramid, where the small size of the images permits an exhaustive search computing the correlation in the Fourier domain. However, the best transformation at any level often is not the one which later will lead to the final solution, because the hierarchic approach introduces false maxima. Therefore, for the sake of robustness we might need to keep several values from each level as seeds for the next. For each level we reduce the number of seeds to the half, to end optimizing a single value for the level of highest resolution.

We have used the downhill Simplex algorithm as implemented in Ref. 11 to maximize the correlation function at all levels except the initial. Each search, started from the seeds from the previous level, finishes when the algorithm achieves a zone where the difference between the maximum and the minimum values found in the neighborhood is lower than a threshold. The tolerance value and number of seeds in the highest resolution level determine to a high degree its robustness and final time.

We have also implemented the optimization search with a simulated annealing algorithm, but it did not improve the results as a rule. Also, we have extended the matching scheme with chamfer distances as described in Ref. 6. This algorithm assigns to each voxel the approximate distance to a surface, i.e., the segmented crease, by means of a fast correlation of the image with two masks. We apply this process only to the first and second levels (higher resolution) of the CT image, and at the following levels the optimization steps run without modification. As a result, the correlation function does not only take into account exact matching of the two surfaces, but also their proximity, and the peak of the function corresponding to the maximum becomes wider and easier to locate.

4 Short Description of the Mutual Information Method

Given a pair of images S and D and a transformation $\mathcal{T}$ to map D into S, voxel-based methods define a measure $M(S_{i,j}, \mathcal{T}(\mathcal{D})_{i,j})$ which is optimized when the images are aligned. M represents a statistical model to relate the value of the pixel i,j from one image to the other. For instance, high intensity values in CT, mostly due to the bone tissue, are likely to be mapped to low intensity values in MR. Since the measure does not make limiting assumptions but builds for each registration a new statistical model, it can potentially use all the information available in both images.

Although early related measures were proposed by Woods, Mazziotta, and Cherry[12] and Hill, Studholme, and Hawkes,[13] the mutual information paradigm was published independently by Maes *et al.*[14] and Wells *et al.*[15] They propose to measure the distance between the joint distribution of two random variables S and D, $p_{SD}(s,d)$, and the distribution associated with their complete independence, $p_S(s) \cdot p_D(d)$. In the image processing field, $p_{SD}(s,d)$ is defined as the normalized number of occurrences of the value s of a pixel $\mathbf{x}$ in S overlapping its corresponding pixel $T(\mathbf{x})$ in D with value d, i.e., the normalized histogram of the overlapping part of S and D.

$$I(S,D) = \sum_{s,d} p_{SD}(s,d) \log \frac{p_{SD}(s,d)}{p_S(s) p_D(d)}, \tag{10}$$

$$I(S,D) = H(S) + H(D) - H(S,D), \tag{11}$$

$$Y(S,D) = \frac{H(S) + H(D)}{H(S,D)}. \tag{12}$$

The original definition of mutual information (10) can be rewritten in terms of entropy (11). Another related measure $Y(S,D)$ is called *normalizing mutual information*, and it is the one we have used in this paper for comparison purposes because it is less sensitive to the size of the area overlapped by the two images. For a complete review of all these measures, see, for instance, Ref. 16.

5 Results

We have applied our method to five pairs of CT-MR images of different characteristics, described in Fig. 6. Pairs 1 and 2 represent the best possible conditions: high resolution, good contrast and wide field of view (Figs. 2 and 3). Pairs 3, 4 and 5 (Figs. 4 and 5) are a challenge because they have a low number of thick slices, their contrast is nonhomogeneous and their MR acquisition settings do not permit an easy segmentation of the bone. At first our method did not converge for data sets 3–5: false creases segmented outside the head in the MR image misled the optimization scheme. The first solution was to increase the number of seeds which are passed from level to level, in order to widen the search space. As a result, the method converged successfully in all the cases, at the price of increasing the time to 40 min. Therefore, we applied a different approach which consisted of removing creases detected outside the head. The mask for filtering was automatically computed in two steps: a simple threshold and a morphological closing. Then, the optimization running on the filtered creases was successful and fast.

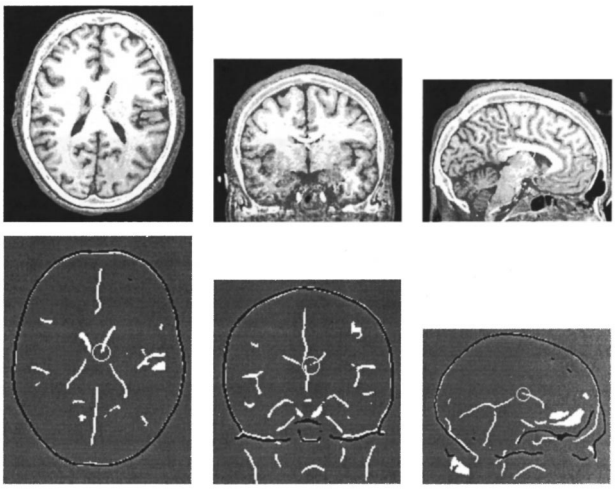


Fig. 3 Fusion of registered volumes of data set 1. From left to right, first row shows MR with CT bone superimposed. Second row shows CT crest (black) registered with MR valley (white). The white circle marks the intersection of the three orthogonal views.

There are two main issues to validate a registration method: accuracy and robustness. The first bounds the error expected when we relate the coordinates of one image into the other, while the second tells the reliability of the method under adverse conditions.

The accuracy can be measured using two approaches. First, we could take as a gold reference the transformation which minimizes the mean square correspondence error of several stereotactic frame landmarks manually pointed out by experts, as in Ref. 3. However, the correspondence error of our transformation is expected to have a similar magnitude to that of the landmarks' approach and needs a much

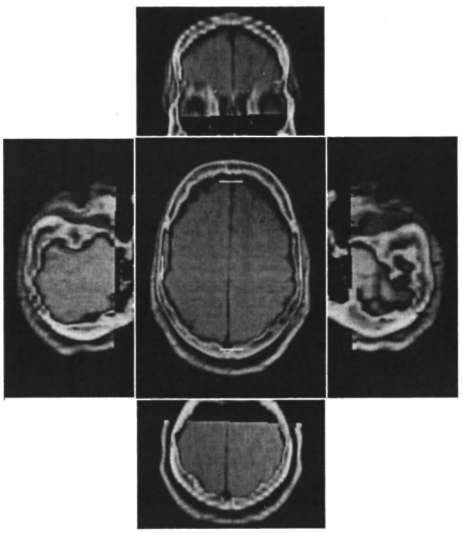


Fig. 4 Fusion of registered volumes of data set 3 showing MR with CT bone superimposed. Note the thin gap with constant width between the border of the bone in the CT and the brain tissue in the MR. The white lines in the central image show the placement of the adjacent cuts.

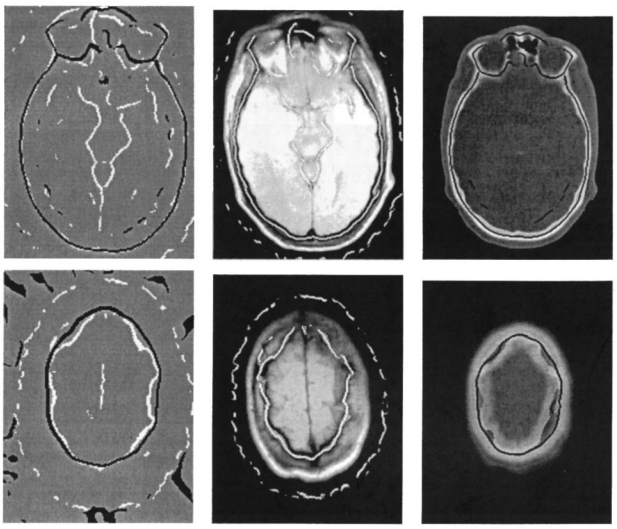


Fig. 5 Registered creases of two planes of data set 3. The left column shows two axial views of corresponding creases from CT (black) and MR (white). The image below shows misregistration because, although the crease follows the valley properly, in the MR modality the upper slices do not depict the bone as a valley but as marrow surrounded by a thin layer of bone. This can be clearly seen comparing the segmentation of the bone in the middle column (in white) for the MR image, and the right column (in black) for the CT image.

larger database of images. Taking this into consideration, we are currently participating in the Vanderbilt project,[2] to assess the overall accuracy of our method when compared to an extrinsic method.

For this experiment, we have chosen to check visually the images fused using the registration results. Figures 3 and 4 show the mix of the CT and the MR images using a point to point maximum operation. In this visualization, we have paid special attention to the inner surface of the bone, which contains a soft tissue, the dura, which appears in black and should be of constant width.

We have visually inspected the registration for all five data sets and found them to be similarly well aligned both for our method and for mutual information. Differences are larger for data sets 3, 4 and 5, and are due to two factors. The first is their low resolution, which magnifies the differences. The second, also important, is that their settings make the two methods converge to two different solutions which actually are the maxima for their respective goodness function. Figure 5 shows several axial slices of the data set 3. The bone is properly segmented for the medium axial cuts, but in higher planes the bone appears filled by marrow, which is visible in the particular settings of these MR modalities but not in the CT image. Therefore, for these two pairs the segmentation done by the creaseness operator is slightly different for each image. However, we think that the registration is still valid because the correlation, which measures the quality of a transformation, takes into account only common structures, which coincide in the rest of the segmentation.

We have carried out a second set of experiments to assess the robustness of the method. The structure of the experiment is as follows: first, we run the described method to register the source images. Next, the CT image is transformed with other known parameters T_T, and the registration algorithm is applied again. An error-free method would return T_R equal to T_T, while a large difference corresponds to misregistration. An important issue is the method to assess the quality of a recovered transformation. We have measured the mean distance between the two coordinates which are the result of transforming *all pixels of the head*, that is, brain and skull, by T_T and T_R. We have built a set of 50 trial transformations with random parameters distributed uniformly with increasing magnitude, from ± 4 to ± 25, both in translations (mm) and rotation angles (deg). See Fig. 9, left columns, for a sample.

Dataset number	Modality	Dimensions x,y	Dimensions z	resolution x,y (mm)	resolution z (mm)	Error Mean (mm)	Error Max (mm)	N<10 (50)	Mean Time (min)	Method
1	CT	256	100	0.93	1.55	0.59	1.47	50	2.61	C
	MR-T1	256	180	0.97	1	2.25	4.72	50	6.78	MI
2	CT	320	128	0.71	1.5	0.35	0.83	49	3.48	C
	MR-T1	256	100	0.9	1.5	1.73	3.61	50	5.97	MI
3	CT	512	29	0.65	4	1.4	3.8	50	3.12	C
	MR-PD	256	26	1.25	4	4.12	9.69	50	3.13	MI
4	CT	512	29	0.65	4	1.92	4.49	40	3.86	C
	MR-T1	256	26	1.25	4	5.34	10.62	50	3.49	MI
5	CT	512	29	0.65	4	1.53	3.86	41	3.46	C
	MR-T2	256	26	1.25	4	6.42	14.11	50	2.97	MI

Fig. 6 Left: specifications of the five pairs of data sets used in our experiments. On the right-hand side, for each data set we compare the global results of the creaseness method (C, first row) and mutual information (MI, second row). The mean and max columns refer to the mean and maximum errors of the 50 trial transformations done for each pair. The next column gives the number of the transformations recovered within a distance of 10 mm from the trial.

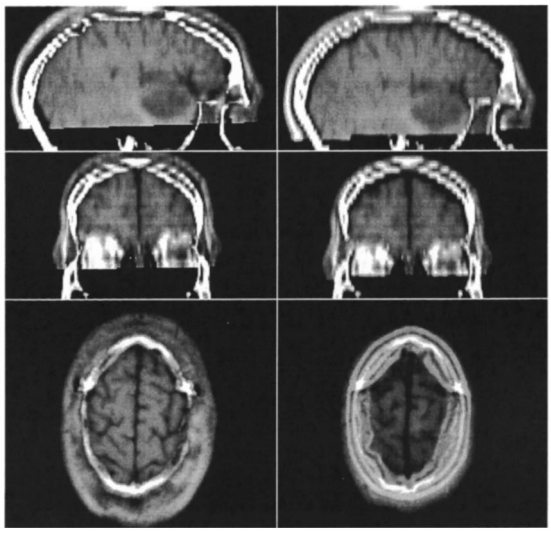


Fig. 7 Data set 4 exhibits the clearest differences between the registration using our method (left) and mutual information (right). These three views (sagittal, coronal and axial) show that it is not easy to decide visually which is better. Our registration may be slightly low along the *z* axis because of the same problems as in data set 3. In contrast, the mutual information transformation seems to position the bone too high, even overlapping the skin on the top.

	A		B		C		D	
Description	Single seed		Multiple seeds		No hierarch.		Chamfer	
Ftol	1e-3		1e-5		1e-5		1e-3	
Seeds per lev.	{1, 1, 1, 1, 1}		{1, 3, 4, 4, 8}		1 -		{1, 3, 4, 4, 8}	
Dataset	1	3	1	3	1	3	1	3
Median (mm)	0.59	1.4	0.54	0.95	0.55	0.99	1.38	1.21
Time (min)	2.61	3.12	8.65	13.15	5.43	7.11	4.31	8.12
# <10 (50)	50	50	49	49	42	42	49	50

Fig. 8 Performance of our method for 50 trial misregistrations and two data sets. Parameters for the search algorithm: Ftol, the tolerance threshold of the optimization method at the highest resolution level, and the number of seeds passed through the hierarchy (from highest to lowest resolution). Method A passes only one seed, and ftol is set relatively low. Its results have not been improved by Method B, despite its more demanding tolerance threshold. Method C shows the necessity of the hierarchical approach: robustness is lower, and even the time to converge is higher because the optimization steps take much longer to run. Method D uses chamfer distances to match the surfaces, with results slightly worse than A.

The first goal was to optimize the parameters in our registration method. There are several main issues to be investigated experimentally: (a) the need for a hierarchical approach, (b) the number of seeds needed at each level, and (c) the tolerance threshold of the optimizing method at the highest resolution level. We have tested three different combinations of parameters, and checked for their robustness and execution time. Figure 8 shows the results for three different choices. It is clear that the hierarchical choice (columns 1 and 2) improved robustness, while keeping the total time low; the time for simple optimization setting (column 3) was approximately double. Although the number of seeds does not seem to improve robustness for this set of misregistration, it has already been successfully employed to improve robustness in the initial registration of datasets 3–5. Matching with chamfer distances (column 4) converges to the solution faster than setting 2, but, on the other hand, it seems to widen the peak of the optimizing function. Since it is not better than setting 1, it may be useful only for cases where corresponding features do not appear (e.g., creases in MR to edges in PET).

The final aim of this paper is to validate the performance of our method against another, which we can take as golden standard. We have chosen the *normalizing mutual information* because it is representative of the voxel-based methods and it has better results reported in other papers.[2] We have set our method to the parameters labeled as ''single seed'' in Fig. 8, and we have run the experiment for 50 trial misregistrations, for five pairs of images. Our algorithm takes 3–5 min to converge on a PC *Pentium* at 350 MHz with 256 Mb of memory running under Linux, plus approximately 5 min$\times$2 images to extract the creases. Results for the mutual information (MI) algorithm were obtained on the same computer and took from 3 to 6 min to complete.

We have distributed the results of this experiment into two figures. Figure 6, right-hand side, provides statistical results for all five pairs of images. Figure 9 gives some of the tried misregistrations and the resulting error for both methods, for data sets 1, 3 and 5 (for all different MR modalities). Our method gives best results for data sets 1 and 2 because the segmentation process is more accurate due to the quality and number of slices of the images. In these sets, it converges more confidently and with a higher reproducibility than the mutual information method. Results from data sets 3, 4 and 5 are less clear: failures are equally high for both methods, and the lower mean of the creaseness method is less significant. The explanation is that the overlapping area of the two original images is narrow, and once the trial transformation T_T has been applied,

Dataset 1: MR-T1

	Trial tranformations						Creaseness							Mutual Information						
N	θ_x	θ_y	θ_z	t_x	t_y	t_z	$\Delta\theta_x$	$\Delta\theta_y$	$\Delta\theta_z$	Δt_x	Δt_y	Δt_z	M.D.	$\Delta\theta_x$	$\Delta\theta_y$	$\Delta\theta_z$	Δt_x	Δt_y	Δt_z	M.D
0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0.72	0.12	0.89	0.12	0.25	0	**1.82**
1	-1.21	-3.01	0.33	3.82	-0.34	0.52	0.35	0.22	0	0.09	0.09	0.01	**0.6**	0.13	0.26	1.16	0.2	0.03	0.11	**1.98**
2	3.54	-1.5	0.66	-1.7	-3.83	-2.2	0.02	0.36	0.08	0.09	0.05	0.03	**0.55**	1.21	0.42	1.2	0.18	0.17	0.08	**2.72**
3	-1.54	-2.03	-2.42	0.04	-1.58	-0.86	0.11	0.44	0.11	0.08	0	0.06	**0.68**	0.17	0.24	0.2	0.16	0.11	0.11	**0.58**
4	-0.69	0.13	0.77	-0.7	0.35	-0.42	0.1	0.1	0.11	0.05	0.02	0.04	**0.27**	0.15	0.11	1.13	0.18	0.1	0.48	**1.97**
45	13.86	7.42	-3.78	15.71	7.04	-5.73	0.1	0.4	0.19	0.06	0.05	0.03	**0.65**	1.21	0.03	0.34	0.09	0.29	0.29	**1.8**
46	-15.81	-13.44	11.78	6.66	-11.71	9.41	0.12	0.24	0.03	0.05	0.01	0.02	**0.38**	0.18	0.2	1.4	0.09	0.16	0.22	**2.25**
47	14.08	12.99	-1.15	-13.41	-12.16	4.03	0.18	0.32	0.1	0.05	0.03	0.02	**0.53**	1.19	0.01	0.52	0.15	0.22	0.09	**1.77**
48	-8.2	-12.44	0.2	-11.2	5.8	12.92	0.01	0.05	0.1	0.06	0.04	0.04	**0.19**	0.15	0.15	0.14	0.05	0.24	0.02	**0.42**
49	11.96	-13.56	10.96	12.36	-7.44	3.4	0.11	0.43	0.19	0.07	0.06	0.08	**0.71**	12.32	0.55	1.18	0.17	0.18	0.1	**2.24**
Mean error for each parameter							0.14	0.34	0.07	0.07	0.05	0.06		0.26	0.29	1.15	0.16	0.16	0.1	

Dataset 3: MR-PD

	Creaseness							Mutual Information						
N	$\Delta\theta_x$	$\Delta\theta_y$	$\Delta\theta_z$	Δt_x	Δt_y	Δt_z	M.D.	$\Delta\theta_x$	$\Delta\theta_y$	$\Delta\theta_z$	Δt_x	Δt_y	Δt_z	M.D.
0	0	0.01	0	0	0	0	**0.01**	1.97	0.42	0.6	0.19	0.08	0	**3.49**
1	0.23	0.19	0.67	0.02	0.02	0	**0.5**	2.76	0.14	1.04	0.24	0.09	2.29	**5.27**
2	1.68	1.42	0.66	0.26	0.1	1.25	**3.93**	1.63	0.54	0.78	0.45	0.54	2.26	**3.99**
3	0.08	0.29	0.06	0.04	0.1	0.34	**0.59**	2.98	0.24	0.32	0.21	0.17	2.68	**5.24**
4	0.37	0.07	0.05	0.01	0	0.67	**0.81**	1.64	0.13	0.95	0.07	0.02	0.42	**3.42**
45	1.47	0.07	0.07	0.15	0.11	1.07	**2.5**	1.03	0.27	0.46	0.04	0.21	2.6	**3.81**
46	0.13	0.31	0.03	0	0.04	0.16	**0.55**	3.52	0.88	1.93	1.22	0.48	3.59	**7.4**
47	0.59	0.45	0.07	0.05	0.39	0.57	**1.29**	1.25	0.23	0.4	0.34	0.16	2.22	**2.8**
48	0.17	0.35	0.05	0.1	0.1	0.43	**0.71**	2.95	0.21	0.8	0.7	0.05	2.58	**5.25**
49	1.11	0.57	0.15	0.18	0.13	0.36	**2.11**	1.16	0.38	0.7	0.64	0.19	2.47	**3.75**
	0.62	0.21	0.07	0.05	0.1	0.64		2.15	0.22	0.71	0.18	0.19	2.43	

Dataset 5: MR-T2

	Creaseness							Mutual Information						
N	$\Delta\theta_x$	$\Delta\theta_y$	$\Delta\theta_z$	Δt_x	Δt_y	Δt_z	M.D.	$\Delta\theta_x$	$\Delta\theta_y$	$\Delta\theta_z$	Δt_x	Δt_y	Δt_z	M.D.
0	0	0	0	0	0.01	0	**0.01**	3.34	0	0.84	0	0.5	0	**5.72**
1	0.35	0.02	0.05	0.01	0.03	0.31	**0.62**	3.71	0.21	1.4	0.18	0.47	0.98	**6.71**
2	0.82	0.35	0.17	0.02	0.11	0.49	**1.52**	0.14	0.54	1.32	0.2	0.08	0.7	**3.21**
3	0.78	0.24	0.07	0.04	0.18	0.16	**1.33**	4.41	0.18	0.44	0.04	0.33	1.11	**7.19**
4	0.89	0.16	0.13	0	0.18	0.47	**1.53**	3.31	0.13	0.95	0.18	0.34	0.58	**5.77**
45	0.66	0.4	0.07	0.07	0.16	0.26	**1.23**	0.94	0.15	0.22	0.15	0.84	0.85	**2.03**
46	1.79	0.12	0.15	0.29	0.56	0.62	**3.03**	4.35	1.12	2.35	0.72	0.21	2.09	**8.47**
47	**	**	**	**	**	**	**	0.24	0.23	0.46	0.03	0.16	0.22	**1.17**
48	1.57	0.22	0.06	0.17	0.45	0.62	**2.64**	4.56	0.03	1.21	0.26	0.51	1.64	**7.52**
49	0.85	0.37	0	0.13	0.21	0.47	**1.54**	0.03	0.16	1.18	0.27	0.32	0.72	**2.7**
	0.83	0.22	0.09	0.04	0.18	0.35		3.75	0.41	0.94	0.11	0.38	1.11	

Fig. 9 First and last five results (for light and heavy misregistration) of the robustness experiment for data sets 1 (above) and 3 and 5 (below). Left column shows trial parameters in mm and deg, while following columns give the absolute difference to the recovered parameters for the two methods. MD is the mean distance error for the voxels within the head; the last row shows the mean of the errors for each parameter and for all transformations applied. Results labeled as ** have an error bigger than 10 mm. Figure 6 gives mean results for all five data sets.

the resulting images may be too distant to converge. Creaseness measure seems to have a maximum in the optimizing function narrower than MI, as can be seen in Fig. 9: for no transformation at all ($N=0$), MI has an error only slightly better than the mean for all the misregistrations, while the creaseness method is almost error free. It is interesting to note that the mean distance (MD) seems independent of the specific tried transformation, but it depends a lot on the quality of the original images. In addition, the last row in this figure suggests that $\Delta\theta_x$ is the largest source of errors compared to the other parameters.

6 Conclusions

We have presented a registration method for CT-MR 3D images based on a new creaseness operator. This operator has the advantage of giving a high response only at significant creases such as the skull in CT and MR images, but not in the background and inner brain, where the degree of anisotropy typical of ridges and valleys is low. We have presented the recovered transformation parameters for a large number of randomly chosen target transformations and we have visually checked the validity of the registration. The optimization by means of the downhill Simplex algorithm has proven to be accurate and fast.

When compared to the mutual information method, our method shows similar robustness to random transformations, with a comparable computing time. Under certain conditions, i.e., small overlapping areas, our method may not be as robust as MI, but, on the other hand, when it converges its reproducibility is higher.

Since our method takes into account only the rigid transformations of the skull and ignores the soft tissues, we think it can suit very well the applications which need to compute only the registration of the skull. For instance, in studies which need to compare images taken over a long time, our method can be useful because the skull is the only undeformable structure of the head, while others are often changed by chronic diseases. A further step could be to apply intensity-based registration algorithms to specific parts of the resulting registered images. This second registration would give a measure of their relative movement and of their volume changes.

Although in this paper we present results only for CT-MR registration, our method has worked successfully for CT and MR intra-modality studies, and we are currently planning to extend it to include registration of PET images. Future work also includes validating the accuracy for a much larger set of images of known transformation parameters.

Acknowledgments

This research has been partially funded by CICYT projects TIC97-1134-C02-02 and TAP96-062 9-C04-03. The authors kindly acknowledge Dr. Petra van den Elsen from Utrecht University for providing two 3D CT and MR data sets and Dr. J. M. Fitzpatrick, head of the project ''Evaluation of Retrospective Image Registration'' (Project No. NIH R01 NS33926-01) from Vanderbilt University for providing us with their image database. We would like to give special thanks to Dr. D. Hill and Dr. D. Hawkes, from the UMDS, Guy's and St. Thomas' Hospital, for allowing access to the mutual information registration software (written by Dr. Collin Studholme). Finally, we thank the reviewers for their useful comments.

References

1. J. Maintz and M. Viergever, ''A survey of medical image registration,'' *Med. Image Anal.* **2**(1), 1–36 (1998).
2. J. West, J. Fitzpatrick, M. Wang, B. Dawant, and C. Maurer *et al.*, ''Comparison and evaluation of retrospective intermodality brain image registration techniques,'' *J. Comput. Assist. Tomogr.* **21**, 554–556 (1997).
3. P. van den Elsen, J. Maintz, E.-J. Pol, and M. Viergever, ''Automatic registration of CT and MR brain images using correlation of geometrical features,'' *IEEE Trans. Med. Imaging* **14**(2), 384–396 (1995).
4. J. Maintz, P. van den Elsen, and M. Viergever, ''Evaluation of ridge seeking operators for multimodality medical image matching,'' *IEEE Trans. Pattern. Anal. Mach. Intell.* **18**(4), 353–365 (1996).
5. J. Maintz, P. van den Elsen, and M. Viergever, ''Comparison of edge-based and ridge-based registration of CT and MR brain images,'' *Med. Image Anal.* **1**, 151–161 (1996).
6. J. H., R. Robb, and K. Holton, ''A new approach to 3-d registration of multimodality medical images by surface matching,'' in *Visualization in Biomedical Computing*, pp. 196–213, SPIE Press, Bellingham, WA (1992).
7. D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach, ''Ridges for image analysis,'' *J. Math. Imaging Vision* **4**, 353–373 (1994).
8. A. López, D. Lloret, and J. Serrat, ''Creaseness measures for CT and MR image registration,'' in *Proceedings of the Conference on Computer Vision and Pattern Recognition'98*, pp. 694–699, IEEE Computer Society, New York (1998).
9. A. López, F. Lumbreras, and J. Serrat, ''Creaseness from level set extrinsic curvature,'' in *Computer Vision-ECCV'98*, Vol. 1407, pp. 156–169, Springer, Berlin (1998).
10. G. Golub and C. van Loan, *Matrix Computations*, Chap. 7.3.1, The Johns Hopkins University Press, Baltimore (1989).
11. W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes*, 2nd ed., Cambridge University Press, Cambridge, UK (1992).
12. R. Woods, J. Mazziotta, and S. Cherry, ''MRI PET registration with automated algorithm,'' *J. Comput. Assist. Tomogr.* **17**, 536–546 (1993).
13. D. Hill, C. Studholme, and D. Hawkes, ''Voxel similarity measures for automated image registration,'' in *Proceedings of Visualization in Biomedical Computing*, R. Ra, Ed., Vol. 2359, pp. 205–216, SPIE, Bellingham, WA (1994).
14. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, ''Multimodality image registration by maximization of mutual information,'' *IEEE Trans. Med. Imaging* **16**(2), 187–198 (1997).
15. W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, ''Multimodal volume registration by maximization of mutual information,'' *Med. Image Anal.* **1**(1), 35–51 (1996).
16. C. Studholme, D. Hill, and D. Hawkes, ''Automated three-dimensional registration of MR and PET brain images by multiresolution optimization of voxel similarity measure,'' *Med. Phys.* **24**(1), 25–35 (1997).

David Lloret received his MS in computer science from the Universitat Autònoma de Barcelona in 1997. He is an assistant professor in the same university and a member of the Computer Vision Center. His research interests are in the area of medical image processing, and especially on registration of CT, MR, and ultrasound images.

Antonio M. López received his BS degree in computer science from the Universitat Politècnica de Catalunya in 1992 and his MS degree in computer science, option image processing and artificial intelligence from the Universitat Autònoma de Barcelona in 1994. Presently, he is an associate professor in the Computer Science Department of the Universitat Autònoma de Barcelona and responsible for the area of development in the Centre de Visió per Computador. He is currently finishing his PhD degree, which is de-

voted to the design and implementation of specific low-level geometrical descriptors with application to medical image analysis problems.

Joan Serrat received his PhD degree in computer science in 1990 from the Universitat Autònoma de Barcelona. Presently, he is an associate professor in the Computer Science Department and a member of the Computer Vision Center at the Universitat Autònoma de Barcelona. His interests are wavelets for texture analysis and low-level geometrical descriptors of images.

Juan J. Villanueva received his BSc in physics from the University of Barcelona in 1973 and his PhD in computer science from the Universitat Autònoma de Barcelona in 1981. Since 1975, he has been teaching at the Computer Science Department, where he was appointed professor in 1990. In 1986, Dr. Villanueva promoted the Computer Vision Center and has been its director since its foundation. He was co-founder and vice-president of AERFAI, which is the Spanish chapter of IAPR; he is currently a member of its steering committee. His research interests comprise all aspects of computer vision, in particular, recognition based on geometrical models and medical computer vision applications. Dr. Villanueva is a member of the IEEE, SPIE, and the IEEE Computer Society.