Rigid registration of CT and MR volumes based on Rothe's creases

David Lloret, Antonio M. López, Joan Serrat
Computer Vision Center, Universitat Autònoma de Barcelona
Edifici 0, 08193 Cerdanyola
david@cvc.uab.es

Abstract

Automatic registration of multimodality images is an upcoming field of research for medical diagnosis and surgical planning. We propose a new method to extract common features from different modalities based on the extension to 3D of Rothe's creases. Extracted images are used as input for a multiresolution search algorithm which brings them into agreement.

Key Words : registration, crease, flowline, multiresolution search.

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 [1]. One of the images has to be transformed geometrically to bring it into agreement with the other one, taken as a reference. Usually, the transform type and/or its parameters are unknown and must be estimated from the images themselves.

The need to register images has arisen recently in the field of medical imaging, with the introduction of 3D modalities such as computed tomography (CT), magnetic resonance (MR), digital x–ray angiography (DSA), positron emission tomography (PET) and single photo emission computed tomography (SPECT). Abnormalities observed in images of functional modalities, like PET and SPECT, can be located more precisely when they are registered with images of anatomical modalities such as CT, and MR, of the same patient [2, 10]. Thus a major application of image registration is in surgery and radiotherapy planning. Several anatomic modalities need to be registered in order to accurately localize a lesion and better understand its 3D spatial relationship with nearby sensitive structures like cerebral vessels [2].

In this paper we propose a method for the automatic registration of 3D CT and MR brain images. It is based on the images themselves instead of relying on external references like skin markers or a stereotactic frame. It aims specifically at intrasubject registration of brain images, thus allowing the assumption that a rigid transform exists between the two volumes. This is an important issue because the space of possible transforms and consequently the search time are substantially reduced. Our method is greatly inspired by the work of van den Elsen and colleagues [10, 6]. They determine the parameters of a rigid transform through the correlation of CT ridge and MR valley images in a hierarchical scheme. In their work, ridges and valleys are computed by means of a generalization to 3D of the classical scale space L_{vv} operator. We propose to change this operator for a better definition and implementation of ridges and valleys, the Rothe's creases, whose study and implementation was described in a previous paper [5]. The aim of this article is to show that Rothe's creases, naturally extended to 3D, are suitable for the problem of CT and MR volume registration and, furthermore, that they exhibit both conceptual and computational advantages over the L_{vv} operator.

This paper is organized as follows. In section 2 we address two major definitions of creases and their extension to 3D. In section 3 we will explain a method for CT and MR volume registration by means of 3D ridge and valley correlation. The next section presents the obtained results. Finally, section 5 discusses conclusions and future work.

2 Creases in 3D

Creases (ridges or valleys) have deserved special attention in geosciences as relevant features of a topographic relief. They are the points where water gathers to run down hill, in the case of valleys, and similarly for ridges with the relief turned upside down. Since the end of the last century researchers have tried to characterize them mathematically. Two main non–equivalent definitions were proposed relying on the modelization of a topographic relief as a 2D height function. The first one, owing to De Saint–Venant, identifies creases as points of extreme slope along a level curve of the relief [8]. This condition can also be found under equivalent formulations: height extrema in the directions where the second directional derivative is also extreme, or loci of curvature extrema along the relief level curves. Therefore, the De Saint–Venant'definition can be implemented locally. The second one, due to R. Rothe, identifies creases as parts of flowlines where other flowlines converge [7]. Flowlines are the integral curves of the gradient vector field of the height function, that is, curves which follow the gradient direction at each point. Thus, a flowline $\mathbf{f}(t)$ lying on the image landscape L can be characterized by the ordinary differential equation $\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}t}(t) = \nabla L(\mathbf{x})$ where t is a given parametrization, $\mathbf{f}(t) = \mathbf{x}$ and ∇L is the image gradient. In fact, Rothe's is the correct solution for the above "hydrologic" definition [3]. This characterization gives rise to semi–local implementations: convergence of flowlines have sense over a given region of interest. Even though the De Saint–Venant definition is not conceptually the right one, it is the most widespread in the computer vision literature, due to the fact of its local character. Nevertheless, this characterization implies second order derivatives to decide whether we are in a crease point. Conversely, according to Rothe's definition, we can take such a decision by calculating just the first order derivatives needed to compute the gradient vector field.

A straightforward approximation of the De Saint–Venant's condition is achieves by selecting locations where the curvature of level curves is high in absolute value. This curvature is a tensorial invariant quantity defined as[1] [9]:

$$\kappa = \frac{L_i L_i L_{jj} - L_i L_j L_{ij}}{(L_k L_k)^{3/2}} \tag{1}$$

In order to avoid some pathological cases [3] where κ fails to detect creases and also to get rid of creases in background areas, κ is multiplied by the gradient magnitude $L_w = \sqrt{L_k L_k} = \parallel \nabla L \parallel$. The resulting expression is the operator $L_{vv} = -L_w \kappa = (L_i L_j L_{ij} - L_i L_i L_{jj})/(L_k L_k)$.

This approximation to the De Saint–Venant's condition has been proposed in the linear scale space context. In it, the ill–posed problem of image derivatives computation is regularized convolving the image with the corresponding sampled gaussian derivative. Thus, for the first order derivatives of a nD image we have $L_\alpha(\mathbf{x};\sigma) = L(\mathbf{x}) * G_\alpha(\mathbf{x};\sigma), \quad \alpha = 1, ..., n$ where $G(\mathbf{x};\sigma) = (2\pi\sigma^2)^{-n/2} \ e^{-\|\mathbf{x}\|^2/2\sigma^2}$. The 2D L_{vv} operator has been applied to multimodality medical image matching in [10] where, to work with 3D MR and CT data, the geometric idea behind this operator was extended to 3D: in order to look for ridges/valleys the authors minimized/maximized a function involving the whole set of 3D second order derivatives at each voxel of the images. This turned out to have a high computational complexity.

In [5] we implemented Rothe's definition to extract ridges and valleys from different types of images (MR, range, intensity, angiography), showing the advantages of this definition in the 2D case. In this paper, we want to investigate the suitability of a 3D extension of Rothe's creases as features for multimodality medical image registration, instead of approximations to the De Saint–Venant's definition. In this way, we will be generalizing the correct characterization of crease points in a landscape. Moreover, this definition involves first order derivatives to obtain crease points and a further minor computation step based on the local gradient behavior to distinguish between ridges and valleys, as we will explain later in this section.

The right ridge/valley characterization, due to Rothe, identifies them as parts of flowlines where other flowlines converge. In the discrete domain we translate the property of *convergence* as *overlapping*.

[1] We use the Einstein summation convention and we denote as L_i the first derivative of L with respect to the i–th axis of any adopted orthonormal coordinate system. Analogously, L_{ij} will be the second order derivatives with respect to the i–th and j–th axes. See [9] for further explanations.

Therefore, the basic idea of our algorithm is as follows: we will trace the flowlines of the 3D landscape, counting how many of them pass through each voxel. At the end of the process we will have a 3D accumulated image which will tell us the degree of convergence or *creaseness* at each voxel.

Once we have selected the desired scale of analysis, σ, we trace each flowline by selecting an unvisited voxel $\mathbf{x}$ and then follow the gradient field in the two opposite directions. Thus, the curve progresses according to

$$\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}s}(s+\delta s) = \mathbf{f}(s) + \frac{\nabla L(\mathbf{f}(s);\sigma)}{\parallel \nabla L(\mathbf{f}(s);\sigma)) \parallel}\delta s \tag{2}$$

where $\mathbf{f}(0) = \mathbf{x}$ and $\delta s = 1$ for one direction and $\delta s = -1$ for the opposite. During this process we need to calculate ∇L not only at points with integer coordinates but also at points in between. However, the expression for $L_\alpha(\mathbf{x};\sigma)$ applies only to integer coordinates. Hence, we approximate subvoxel derivatives by a trilinear interpolation of them at the eight nearest voxels. There are three reasons to stop the process of curve tracing in a given direction: 1) the curve reaches a border of the image, 2) a critical point ($\parallel \nabla L \parallel = 0$), or 3) it starts to jump forward and backward from one side to the other of a singular line of the gradient (among them, ridges and valleys). This last event can be checked by comparing the gradient direction at consecutive points.

To trace curves from all the voxels of the image, would be very time consuming. Instead, we assume that once a flowline has passed through a voxel, another one starting from that voxel would be the same flowline. This is an interpretation of the uniqueness theorem for ordinary differential equations, assuming that a point of an image curve has the volume of a voxel. Hence, we limit the number of computed curves to those which are needed in order to cover the whole 3D image, making sure that at least one curve passes through each voxel. Experiments showed that we can obtain almost the same creaseness information by this procedure as by tracing curves starting from all the voxels.

Once the creaseness image has been computed, we have to decide whether the creaseness is ridgeness or valleyness, which depends on the convexity or concavity of the image with respect to its intensity axis. This information is provided by the sign of κ: if $\kappa > 0$ the voxel is classified as ridge, otherwise as valley. For the sake of efficiency, we compute κ by a *literal* implementation of the following equation:

$$\kappa(\mathbf{x};\sigma) = -\mathbf{div}\Big(\frac{\nabla L(\mathbf{x};\sigma)}{\parallel \nabla L(\mathbf{x};\sigma) \parallel}\Big) = -\sum_{i=1}^{3} \frac{\partial}{\partial x_i}\frac{L_i(\mathbf{x};\sigma)}{\parallel \nabla L(\mathbf{x};\sigma) \parallel} \tag{3}$$

which turns out to be more efficient in time and memory [4] than its tensorial counterpart (1). This is due to the fact that equation (3) avoids the use of second order cross–derivatives L_{12}, L_{13}, L_{23} by computing the divergence of the normalized gradient. The derivatives of the normalized gradient in equation 3 are approximated by centered finite differences.

3 Registration method

Any registration method has to specify several issues [1] : 1) the image features from which the transformation will be estimated (e.g. contours, intensities, characteristic points), 2) a similarity measure between the reference image and the transformed one, assessing the "goodness" of that particular transformation, and 3) the search space and search strategy, namely, the space of allowed transforms and how to reach within it the location where the similarity measure is maximum.

The skull is a structure visible both in CT and MR brain images. 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 strong signal of the skull in CT forms a ridge, whereas in MR it becomes a valley (figure 1). Thus, pairs of images displaying those features are suitable for the problem of CT and MR registration. In particular, we intend to find out the transformation parameters which better align CT ridgeness with MR valleyness accumulated images. We have chosen as a similarity measure the cross correlation of the former creases images. Thus, the correlation will give a maximum when the search in the parameter space leads us to a transformation which better aligns the images.

A rigid transformation models with reasonable accuracy the geometric difference between CT and MR brain images because the skull does not deform . Rigid transformations are characterized by three rotation, translation and scaling parameters, which form a 9-dimensional search space. The scaling parameters, however, are usually known because they can be deduced from the resolution of the acquisition devices. In 3D, an exhaustive search is an outer possibility due to the high dimensionality of the search space and because it implies executing two time-demanding processes for each point within it: transformation and correlation. Therefore, a partial or heuristic search must be carried out. This poses a second problem, the possibility of getting trapped into a local instead of global correlation maximum. An approach to overcome these two problems is to search the parameter space at multiple resolutions. We handle multiple resolution by building two pyramids where the original CT and MR images are at the bottom and each level is a sampled version of the previous at half resolution. We sample each image by replacing every 8 neighbor voxels by a single one, the maximum of their values. This sampling has the advantage to prioritize interesting intensity features before empty useless zones. In our algorithm, the size of the bottom image is $128 \times 128 \times 128$ while the size of the top is $32 \times 32 \times 32$, which is small enough to be useful and still contains meaningful information.

As a general rule, the search for the best transformation at each level is performed around the results found in the previous one. We produce all the combinations of 5 steps around each parameter (step size 0.05 rad. for rotations and 1 pixel for translations) and score them using the cross correlation. The scores are merged to the list of the 25 best of the level, which will become the seeds for the next one. The only exception is the top level, where the search is done exhaustively to take advantage of its reduced dimensions ($32 \times 32 \times 32$). The output result is the best score from the list produced from the bottom level ($128 \times 128 \times 128$).

4 Results

Original images were kindly provided by Dr. Petra van den Elsen. The size of MR image was $256 \times 256 \times 256$ voxels with resolution $0.9765 \times 0.9765 \times 1.0$ mm/voxel and the CT image was $256 \times 256 \times 100$, with resolution $0.9375 \times 0.9375 \times 1.55$mm/voxel. In order to perform registration experiments, the CT image was transformed to match the MR, and then both were scaled at half resolution to speed up computations. Figure 2 shows several orthogonal planes of the original and accumulated images of MR and CT. Many hidden features now become apparent, namely the skull in both modalities and brain convolutions in MR. Note that this accumulated image includes both ridgeness and valleyness. The input images of the registration algorithm are shown in figure 3. We can see that valleys of MR closely match crests of CT within the area of the skull.

The general idea of the experiment is as follows: given a CT and a MR image already brought into agreement, we transform the CT with certain target parameters θ_x,θ_y,θ_z and t_x,t_y,t_z.Then, we try to recover these parameters by registering the creases of the transformed CT with the valleys of the original MR. If we are able to do so, we presume that we can reliably bring the two original images into agreement.

A perfect registration should recover parameters equal to the input target parameters. Table 1 shows some preliminary results of our experiments. The accuracy is usually good for rotation parameters, but not as good for translation. Extracting creases from an image takes about 5 minutes, i.e., 2 minutes for the completion of derivatives L_1, L_2, L_3, 2 for the accumulation and 1 for the ridge/valley selection. The parameter search algorithm takes much more time, about 3 hours, to complete. Programs have been executed in an IBM SP2 with 128MB RAM.

Target transformation						Recovered transformation					
θ_x	θ_y	θ_z	t_x	t_y	t_z	θ_x	θ_y	θ_z	t_x	t_y	t_z
0.0	0.0	0.45	-8.0	5.0	0.0	0.0	0.0	0.45	8.0	6.0	0.0
-0.5	0.3	0.2	3.0	1.0	-8.0	-0.5	0.3	0.2	3.0	1.0	-10.0
0.3	-0.3	0.3	-5.0	3.0	-2.0	0.3	-0.3	-0.25	-3.5	0.5	-1.5
-0.2	0.0	0.1	7.0	4.0	-1.0	-0.25	-0.05	0.05	8.0	4.0	-3.0

Table 1: Results of the registration algorithm. Angles are in radians and translation in pixels

The final aim of our work is to bring into agreement two images of CT and MR, in order to display their interesting features together.We have focused only on the selection of image features and a similarity measure. We haven't, as yet, devoted much attention to the optimization of the search strategy.

5 Conclusions and future work

The results found so far suggest that our definition of creases is useful for registration purposes. Furthermore, creases are extracted quickly and clearly show some features which seem to be lacking in some other definitions. We believe them to be related to the brain convolutions and worthy of exploration. However, the search algorithm is still slow and not very accurate. Clearly, it needs further improvement. Future work will include the design of better adapted algorithms. Several possibilities are decoupling the rotation and translation parameters in order to narrow the search space, genetic algorithms and surface-based registration methods.

Acknowledgments

This work has been supported by grant TAP96–0629–c04–03 of the Comisión Interministerial de Ciencia y Tecnología (CICYT). We want to thank the Centre de Supercomputació de Catalunya for allowing us to work with their facilities. The authors gratefully acknowledge also Dr. Petra van den Elsen for providing us with two CT and MR volume datasets used for this research.

References

[1] L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–375, 1992.

[2] D.L. Hill. *Combination of 3D medical images from multiple modalities*. PhD thesis, Unit of Medicine and Dental School (UMDS), University of London, December 1993.

[3] J.J. Koenderink and A.J. van Doorn. Local features of smooth shapes: ridges and courses. In *Geometric Methods in Computer Vision II*, volume 2031, pages 2–13. SPIE, 1993.

[4] A.M. López, F. Lumbreras, and J. Serrat. Efficient computation of local creaseness. Technical Report 15, Computer Vision Center, 1996.

[5] A.M. López and J. Serrat. Tracing crease curves by solving a system of differential equations. In B. Buxton and R. Cipolla, editors, *IV European Conference on Computer Vision*, number 1064 in Lecture notes in Computer Vision, pages 241–250, 1996.

[6] J.B.A. Maintz, P.A. van den Elsen, and M. Viergever. Evaluation of ridge seeking operators for multimodality medical image matching. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 18(4):353–365, April 1996.

[7] R. Rothe. Zum problem des talwegs. *Sitzungsber. Berliner Math. Gesellschaft*, (14):51–69, 1915.

[8] De Saint-Venant. Surfaces à plus grande pente constitutées sur des lignes courbes. *Bulletin de la soc. philomath. de Paris*, March 6th 1852.

[9] B.M. ter Haar Romeny, L.M.J. Florack, J.J. Koenderink, et al. Cartesian differential invariants in scale–space. *Journal of Mathematical Imaging and Vision*, 3:327–348, 1993.

[10] P.A. van den Elsen, J.B.A. Maintz, E.D. Pol, et al. Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Trans. on Medical Imaging*, 14(2):384–396, June 1995.

Figure 1: a) sagital plane of MR b) same plane of CT c) CT skull in white on the MR plane

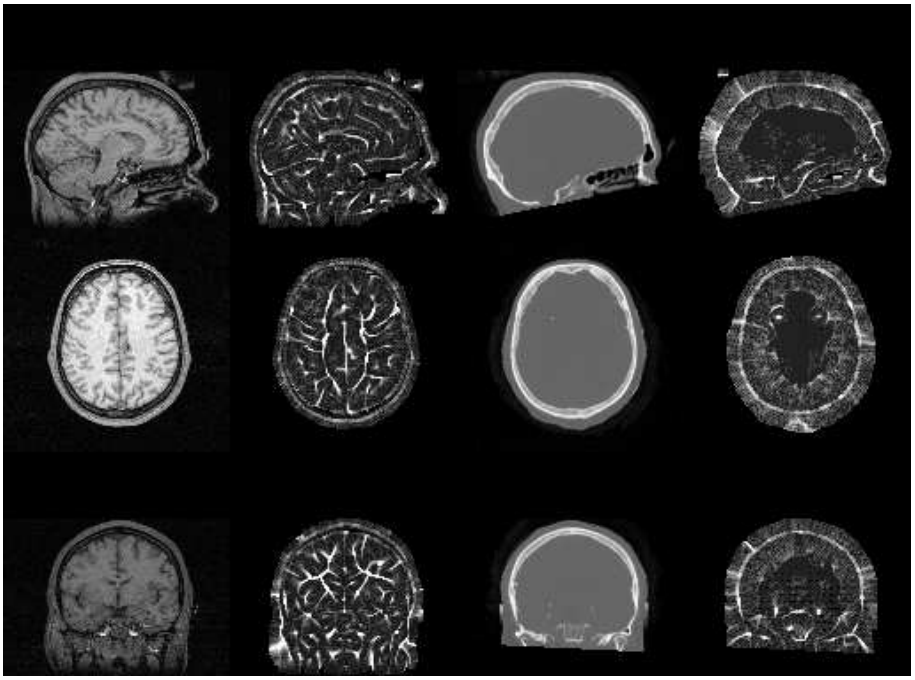

Figure 2: Rows 1-3 show sagital, axial and coronal views. From left to right, each column displays a) the original MR, b) accumulation image of MR c) original CT d) accumulation image of CT

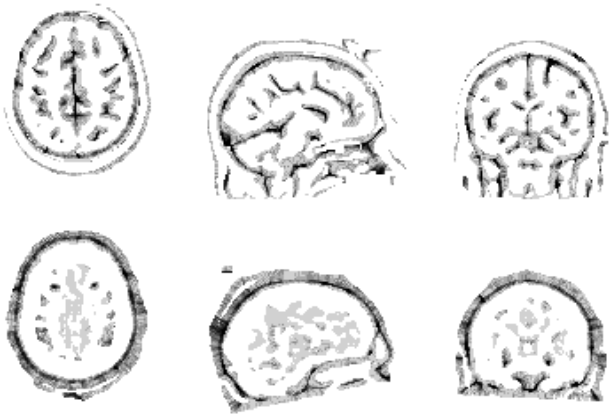

Figure 3: above: MR valleys, below: CT ridges