Block Matching: a General Framework to Improve Robustness of Rigid Registration of Medical Images

S. Ourselin, A. Roche, S. Prima, and N. Ayache

INRIA Sophia - Epidaure Project
2004 Route des Lucioles BP 93
06902 Sophia Antipolis Cedex, France
`{Sebastien.Ourselin, Alexis.Roche, Sylvain.Prima,`
`Nicholas.Ayache}@sophia.inria.fr`

Abstract. In order to improve the robustness of rigid registration algorithms in various medical imaging problems, we propose in this article a general framework built on block matching strategies. This framework combines two stages in a multi-scale hierarchy. The first stage consists in finding for each block (or subregion) of the first image, the most similar subregion in the other image, using a similarity criterion which depends on the nature of the images. The second stage consists in finding the global rigid transformation which best explains most of these local correspondances. This is done with a robust procedure which allows up to 50% of false matches. We show that this approach, besides its simplicity, provides a robust and efficient way to rigidly register images in various situations. This includes for instance the alignment of 2D histological sections for the 3D reconstructions of trimmed organs and tissues, the automatic computation of the mid-sagittal plane in multimodal 3D images of the brain, and the multimodal registration of 3D CT and MR images of the brain. A quantitative evaluation of the results is provided for this last example, as well as a comparison with the classical approaches involving the minimization of a global measure of similarity based on Mutual Information or the Correlation Ratio. This shows a significant improvement of the robustness, for a comparable final accuracy. Although slightly more expensive in terms of computational requirements, the proposed approach can easily be implemented on a parallel architecture, which opens potentialities for real time applications using a large number of processors.

1   Introduction

Numerous intensity-based methods have been successfully devised for rigid registration of medical images [7]. For the monomodal as well as the multimodal case, the general approach consists in assuming a global relationship between the intensities of the images to register (e.g., affine, functional, statistical, etc.) [12],

and then deriving and maximizing a suitable similarity measure (resp. correlation coefficient [1], correlation ratio [13], mutual information [16, 6], etc.). Several issues are common to these methods.

First, all the similarity measures are known to be highly non-convex with respect to the transformation parameters. Thus, their global maximization is seldom straightforward. In most implementations, for sake of low computation time, this latter is performed using a standard deterministic optimization technique (simplex or Powell's method, gradient descent, etc.) which is easily trapped in local maxima. Moreover, the non-convexity of the criterion is enhanced by the interpolation procedure which is used when considering sub-voxel transformations [10].

Second, the assumption of a global relationship between the image intensities may be violated by the presence of various image artefacts. For instance, intensity inhomogeneities are found in several image modalities (bias field in MR, shadowing in ultrasound). Also, biological changes occur in temporal studies (evolutive lesions, tumors, ischemia, etc.) and, more generally, the anatomical structures to be matched may not be strictly equivalent (alignment of serial sections, mid-sagittal plane computation, etc.). These artefacts or anatomical disparities can bias a global criterion and hamper its maximization.

In this paper, we address these problems using a block matching strategy interleaved with a robust transformation estimator. Block matching techniques have already been used in non-rigid medical image registration [2, 4, 8], but very rarely in rigid registration. Instead of maximizing a global similarity measure with respect to the transformation parameters, they treat the problem in a sequential manner. The first step is to build a displacement field by locally translating image regions (or blocks) so as to maximize a given similarity measure; thus, the global maximization is replaced by several *local* maximizations. The second step is to determine the spatial transformation by regularizing the previous displacement field. This two-step procedure is generally applied iteratively in a multi-scale hierarchy.

When using this approach to estimate a rigid transformation, the major difficulty is that the displacements to be found are generally much larger than in a non-rigid context. Therefore, the blocks have to be moved in wider neighborhoods, which may cause the resulting displacement field to contain some severe outliers. These outliers may strongly affect the estimation of the rigid transformation. For that reason, the robustness of the transformation estimator is a key issue.

Our implementation of block matching in rigid mode is described in Section 2, while Section 3 presents results obtained in three different rigid registration problems: alignment of histological sections, computation of the mid-sagittal plane in brain images, and multimodal registration of CT with MR images.

2   Description of the method

The algorithm takes two images as input: a reference image I and a floating image J. The output will be the transformation T and the image $J' = J \circ T^{-1}$, which is aligned with I. The whole process follows from an iterative scheme where, at each step, two successive tasks are performed. The first is computing a displacement field between I and the current floating image J'; this is done through a block matching strategy. The second is gathering these displacements to determine a rigid transformation S using a robust estimator. Updating the current transformation according to $T \leftarrow S \circ T$, we get the new floating image J' by resampling only once the image J in terms of the new T. This two-step procedure is integrated within an iterative multi-scale scheme.

In this section, we have chosen to describe the 3D implementation of the method. However, it should be clear to the reader that the 2D adaptation is straightforward.

2.1   Computation of correspondences by a block matching strategy

The basic principle is to move a block $\mathcal{A}$ of the floating image in a neighborhood Ω, and to compare it to the blocks $\mathcal{B}$ that have coincident positions in the reference image (see Figure 1). The block size and the size of the search neighborhood are constant at a given scale level; they will be respectively denoted $N_x \times N_y \times N_z$ and $\Omega_x \times \Omega_y \times \Omega_z$. The best corresponding block $\mathcal{B}$, with respect to a given similarity measure, allows one to define a match (a_i, b_i) between the centers of the blocks $\mathcal{A}$ and $\mathcal{B}$. In order to improve the computation time, we subsample the search set Ω by considering tentative matches every Σ_x (resp. Σ_y, Σ_z) voxels along the x (resp. y, z) direction. These parameters determine the *resolution* of the displacement field. Also, the centers of the test blocks $\mathcal{A}$ are taken from a sub-lattice of the floating image. We thus introduce another set of subsampling factors $(\Delta_x, \Delta_y, \Delta_z)$, which determines the *density* of the displacement field.

Since the $(N, \Omega, \Sigma, \Delta)$ parameters are voxel distances they are always integers. To take into account possible image anisotropy, their relative values are chosen so as to correspond roughly to the same *real* distances.

Contrary to methods based on a global criterion, the block matching furnishes explicit point correspondences. However, we should suspect some of them to be unreliable. Bad matches often occur when the displaced block is uniform or when it includes structures that are absent from the reference image. To handle outliers, we first reject matches corresponding to similarity measures that are below a given threshold. In this manner, a primary set of outliers is detected. The remaining outliers will be rejected during the robust estimation of the transformation parameters (see Section 2.2).

Multi-scale scheme When choosing large search areas Ω, we expect the algorithm to recover gross displacements while accurate matches may be achieved by assigning low values to Δ and Σ. Moreover, the complexity of our block matching

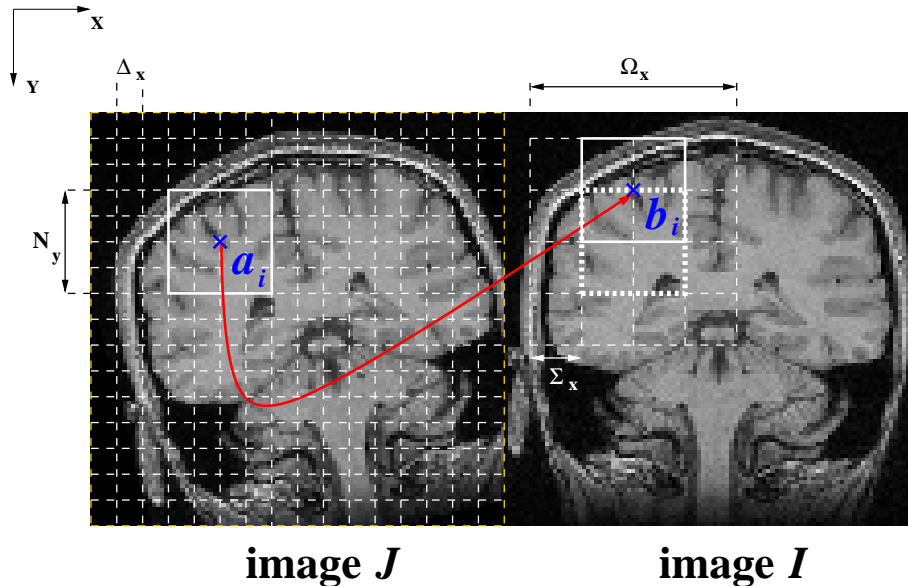**Fig. 1.** *Illustration of the block matching strategy. In a given direction, the parameters have the following meaning: N is the block size; Ω is the size of the search set; Σ is the resolution of the displacement field; Δ is the density of the displacement field.*

procedure is proportional to $(N \times \Omega)/(\Delta \times \Sigma)$. To optimize the trade-off between performance and computational cost, we make use of a multi-scale scheme. We start out with a coarse scale, large values for N, Ω, Δ, and Σ. When refining the scale, all of these parameters are halved; in this manner, the complexity is independent of the scale level. The details of this multi-scale implementation will be found in [9].

Choice of the similarity measure We have implemented several similarity measures. In practice, the correlation coefficient (CC) is well suited for monomodal registration, whereas the correlation ratio (CR) and the mutual information (MI) are better adapted to multimodal registration.

An important fact to bear in mind is that the cost of the block matching procedure is strongly dependent on the time required to evaluate the similarity measure between two blocks. In particular, the MI measure yields computation times that are often prohibitive. Contrary to the case of CR, computing MI between two blocks requires first evaluating their joint histogram [13]. Then, one still has to perform a considerable number of operations, mainly because $n_a \times n_b$ logarithm computations are needed, n_a and n_b being the respective number of intensity classes in the blocks $\mathcal{A}$ and $\mathcal{B}$ to be compared. As a consequence, the computational cost of the block matching procedure using MI is one order of magnitude greater than when using CR.

2.2 The least trimmed squares minimization

A least trimmed squares (LTS) regression is used to find the rigid transformation T that best fits the displacement field computed through block matching. This minimization scheme has been proven to be far more robust to outliers than the classical least squares (LS) method [15]. It consists in solving the following minimization problem,

$$\min_T \sum_{i=1}^{h} \|r_{i:n}\|^2, \tag{1}$$

where $\|r_{i:n}\|^2$ are the squared ordered Euclidean norms of the residuals, $r_i = a_i - T(b_i)$, n is the total number of displacement vectors, and h is set to $h = \lfloor n/2 \rfloor$ to achieve a 50% breakdown point.

As discussed in [14], the LTS criterion may be minimized by means of a simple iterative LS estimation. This technique, generally requiring no more than 10 iterations, has the advantage of being computationally very efficient owing to the fact that the solution of one LS minimization is analytical [3]. We notice, however, that this is not guaranteed to converge to the global minimum of the LTS criterion, but only to a local minimum. Although we have not yet observed cases where the iterative LS was trapped in aberrant solutions, more robust LTS minimization strategies [14] will be looked into in future implementations.

3 Results and discussion

We have applied this block matching algotrithm to three different medical image registration problems. First, we present the histological slices alignment [9]. Second, we present the computation of the mid-sagittal plane [11]. Finally, we have evaluated the robustness of our method with the CT/MR registration problem.

3.1 Histological slices alignment

A similar block matching scheme was applied to estimate a 2D rigid transformation between successive anatomical brain slices [9] (see Figure 2 left). In particular, we realigned several datasets of rat's brains containing about 25 sections each with a resolution of 768×576 pixels ($0.03mm \times 0.03mm$) and an intersection gap of $0.4mm$. The maximization of the global CC [5] were also applied. Our method yields the best visual results, in particular for the inner strucures (see Figure 2 right).

3.2 Mid-sagittal plane computation

We present in Figure 3 the computation of the mid-sagittal plane for MR and SPECT images. The MR image has a size of 256^3, with a voxel size of $0.78mm^3$. The SPECT image has a size of 64^3, with a voxel size of $3.12mm^3$. In [11], an analysis of the accuracy and the robustness of our method has been presented
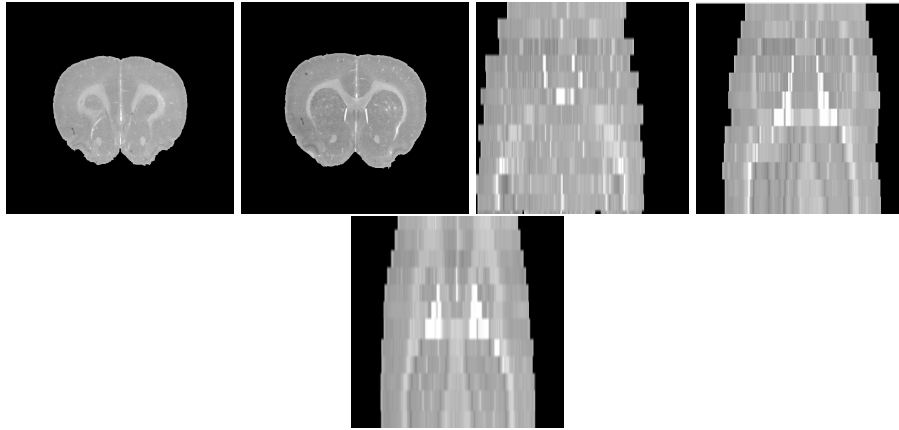
Fig. 2. *Two successive slices and the median perpendicular view of the section set. Middle, initial data. Middle right, after registration using global CC. Right, after registration using the block matching.*

using artificial and real images. Several image databases have been created in order to evaluate the influence of anatomical differences. A number of experiments have demonstrated the robustness and the subvoxel accuracy of our algorithm.
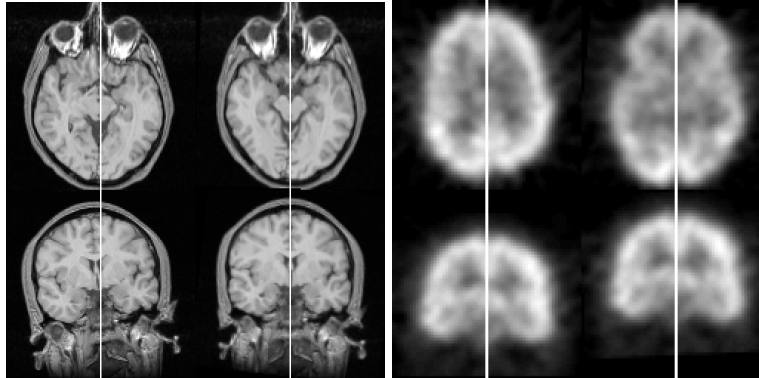

Fig. 3. *axial (top) and coronal (bottom) views, for the initial MR and SPECT and the reoriented and recentered image.*

3.3 CT/MR rigid registration

We have applied our block matching algorithm to the 3D registration of CT with MR (T1 weighted) images. We used eight brain image pairs from the Vanderbilt database (see Figure 5), each ground-truth transformation being known thanks to a prospective, marker-based method [17]. As we are faced with a multimodal registration problem, we have chosen CR as the similarity measure to drive the block matching. Another choice, such as MI, is theoretically possible but, as discussed in Section 2.1, this is prohibitive in terms of computation time.

In these experiments, the block matching algorithm is compared with two methods based on the maximization of a global similarity measure, respectively CR and MI. In these cases, the CR or MI measure was computed using partial volume interpolation (PV) [6], and the maximization was undertaken using Powell's method in a two-step multiresolution scheme (the CT was subsampled by factors $4 \times 4 \times 1$ in the first step, and not subsampled at the second).

In all the ($3 \times 8 = 24$) registration experiments, the initial transformation between the CT and the MR was set to the identity. The registration errors we obtained with the different methods are represented in Figure 4 and Table 1. On the one hand, we consider the *center error*, i.e., the error made on the translation of the CT image center. However, since this error does not reflect the errors in the estimation of the rotation, we also define the *corner error* as the RMS error computed over the eight vertices of a bounding box centered in the CT image, with a relative size of $33\% \times 33\% \times 100\%$. This box is roughly tangential to the patient's head. Of course, to be meaningful, these errors have to be compared with the images resolution: $1.25 \times 1.25 \times 4$ mm^3 for the MR, and $0.65 \times 0.65 \times 4$ mm^3 for the CT.

Table 1. Mean and median registration errors found using three different algorithms in CT/MR registration.

Method	Center errors		Corner errors	
	Mean (mm)	Median (mm)	Mean (mm)	Median (mm)
Global CR	6.57	6.29	7.45	5.80
Global MI	5.08	3.11	5.07	3.08
Block Matching with CR	2.39	2.03	3.70	3.32

These comparative performances lead us to the following observations. First, it is clear that the global methods have both failed in one case (actually, for the same patient dataset). Notice, it was possible to obtain better results by manually resetting the initial transformation parameters, but these are not presented here since our goal is to compare fully automatic methods. We warn the reader that some other results obtained with the global methods might correspond to a local, but non-global, maximum of the similarity measure.

As may be observed in Table 1, the block matching algorithm has provided the best average and median center errors. However, the trend is different when
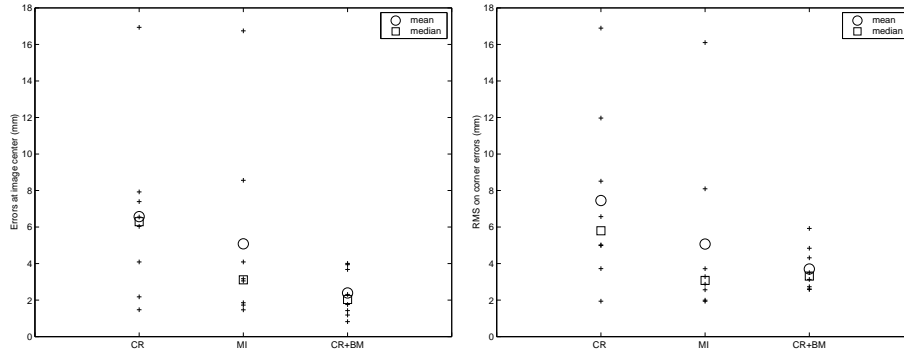
Fig. 4. Registration errors of three different algorithms in CT/MR registration: CR (global maximization), MI (global maximization), and CR+BM (the present block matching algorithm using CR as local similarity measure).

looking at the corner errors, as the performance of global MI and block matching are actually comparable (better for the average for the block matching, while better for the median for global MI). This means that the rotation was best estimated when using global MI: mean rotation errors are 2.85° (global CR), 0.75° (global MI), and 1.68° (block matching).

Another observation is that the performances of the block matching algorithm are remarkably stable across experiments. This is well reflected in Figure 4, where one sees that the error measures corresponding to the block matching algorithm are more concentrated than those corresponding to global CR and global MI. We may thus conclude that the block matching algorithm has provided more robust, if not accurate, estimates of the actual transformations than the global approaches based on MI and CR.
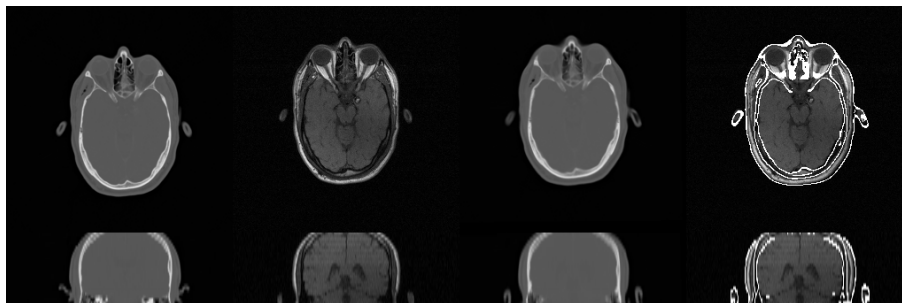

Fig. 5. *CT/MR registration. Left: original CT; Center left: original MR; Center right: registered CT to MR; Right: MR with registered CT contours superimposed*

4 Conclusion

In this paper, we have presented a general strategy for the rigid registration of medical images which is based on the combination of a block matching technique with a robust transformation estimator. We have presented experiments of histological slices alignment, computation of the mid-sagittal plane in brain images, and multimodal CT/MR registration. Our results suggest that, in some cases, the block matching approach can significantly outperform global approaches in terms of robustness.

Concerning the first two experiments, previous analysis [9, 11] has demonstrated the accuracy and robustness of our method. Further analysis is needed in order to better demonstrate its interest in other multimodal registration problems. Moreover, a comparative study between mutual information and correlation ratio within the block matching approach will be investigated.

5 Aknowledgments

References

1. G.L. Brown. A Survey of Image Registration Techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.
2. D. L. Collins and A. C. Evans. ANIMAL: Validation and Applications of Nonlinear Registration-Based Segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 11(8):1271–1294, 1997.
3. D.W. Eggert, A. Lorusso, and R.B. Fisher. Estimating 3D Rigid Body transformations: A Comparison of Four Major Algorithms. *Machine Vision Applications, Special Issue on Performance Characteristics of Vision Algorithms*, 9(5/6):272–290, 1997.
4. T. Gaens, F. Maes, D. Vandermeulen, and P. Suetens. Non-rigid Multimodal Image Registration Using Mutual Information. In *Proc. MICCAI'98*, volume 1496 of *Lecture Notes in Computer Science*, pages 1099–1106, Cambridge, Massachusetts (USA), October 1998. Springer.
5. L.S. Hibbard and R.A. Hawkings. Objective image alignment for three-dimensional reconstruction of digital autoradiograms. *Journal of Neuroscience Method*, 26:55–74, 1988.

6. 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.

7. J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

8. J.B. A. Maintz, E.H.W. Meijering, and M.A. Viergever. General multimodal elastic registration based on mutual information. In K. M. Hanson, editor, *Medical Imaging: Image Processing*, volume 3338 of *Proc SPIE*, Bellingham, WA, 1998. SPIE Press.

9. S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache. Reconstructing a 3D Structure from Serial Histological Sections. *Image and Vision Computing*, 2000. In press.

10. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual information matching and interpolation artefacts. In K. M. Hanson, editor, *Medical Imaging: Image Processing*, volume 3661 of *Proc SPIE*, Bellingham, WA, 1999. SPIE Press.

11. S. Prima, S. Ourselin, and N. Ayache. Computation of the Mid-Sagittal Plane in 3D Images of the Brain. In *Sixth European Conference on Computer Vision, ECCV'2000*, Lecture Notes in Computer Science, Dublin, Ireland, June 2000. Springer.

12. A. Roche, G. Malandain, N. Ayache, and S. Prima. Towards a Better Comprehension of Similarity Measures used in Medical Image Registration. In *Proc. MICCAI'99*, volume 1679 of *LNCS*, pages 555–566, Cambridge (UK), October 1999.

13. A. Roche, G. Malandain, X. Pennec, and N. Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. In *First International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 1496 of *Lecture Notes in Computer Science*, pages 1115–1124, Cambridge (USA), October 1998. Springer.

14. P.J. Rousseeuw and K. Van Driessen. Computing LTS Regression for Large Data Sets. Technical report, Statistics Group, University of Antwerp, 1999.

15. Peter J. Rousseuw and Annick M. Leroy. *Robust Regression and Outlier Detection*. Wiley Series in Probability and Mathematical Statistics, first edition, 1987.

16. P. Viola. Alignment by Maximisation of Mutual Information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

17. J. West and al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Comp. Assist. Tomography*, 21:554–566, 1997.