BLOCK-MATCHING STRATEGIES FOR RIGID REGISTRATION OF MULTIMODAL MEDICAL IMAGES

Olivier Commowick[⋆] *Nicolas Wiest-Daesslé*[†] *Sylvain Prima*[⋆]

[⋆] INRIA, INSERM, VisAGeS U746 Unit/Project, F-35042 Rennes, France
University of Rennes I-CNRS UMR 6074, F-35042 Rennes, France
[†] CHU, University Hospital of Rennes, F-35043 Rennes, France

ABSTRACT

We propose and evaluate a new block-matching strategy for rigid-body registration of multimodal or multisequence medical images. The classical algorithm first matches points of both images by maximizing the iconic similarity of blocks of voxels around them, then estimates the rigid-body transformation best superposing these matched pairs of points, and iterates these two steps until convergence. In this formulation, only discrete translations are investigated in the block-matching step, which is likely to cause several problems, most notably a difficulty to tackle large rotations and to recover subvoxel transformations. We propose a solution to these two problems by replacing the original, computationally expensive, exhaustive search over translations by a more efficient optimization over rigid-body transformations. The optimal global transformation is then computed based on these local blockwise rigid-body transformations, and these two steps are iterated until convergence. We evaluate the accuracy, robustness, capture range and run time of this new block-matching algorithm on both synthetic and real MRI and PET data, demonstrating faster and better registration than the translation-based block-matching algorithm.

Index Terms— MRI, multimodal rigid-body registration, optimization, block-matching.

1. INTRODUCTION

Pairwise multimodal/multisequence rigid-body registration is ubiquitous in medical image analysis [1, 2]. The problem arises when it comes to compare images of the same patient obtained using different modalities (typically, MRI vs CT or PET), or using MRI only but with different pulse sequences (typically, T1- vs T2-weighted). Hereafter, for the sake of clarity, we use the same term "multimodal" to refer to either different modalities or different MRI sequences. Due to the different contrasts of the images to register in this context, the first successful methods used features extracted from each of the two images (e.g. points, curves, surfaces) and then matched to register these images [3]. Since the introduction of the first multimodal similarity measures in the mid-90s (mutual information) [4, 5], intensity-based (iconic) registration methods have somewhat superseded feature-based techniques, mostly due to their higher genericness: on one part, it can prove difficult to obtain homologous features between some image modalities (e.g. MRI vs PET), while on the other part the same similarity measure can potentially be used for any type of multimodal registration.

An important class of intensity-based methods uses similarity measures within a block-matching setting. This strategy was first proposed in the context of motion estimation for video compression [6] and has been subsequently applied to medical image processing [7]. Within this setting, point-to-point correspondences between the two images are first found, based on the intensity similarity of blocks of voxels centered on these points, before the global rigid-body transformation best superposing these matched points is computed. This two-step process (local block-matching and global transformation estimation) is then iterated until convergence. Hereafter, for the sake of conciseness and clarity, we name "block-matching algorithm" this whole iterative process, although this name only refers to its first step, strictly speaking.

Such block-matching algorithms are very powerful, as they benefit from the good properties of both intensity-based (genericness, multimodality, ability to use all the data points for **accuracy**) and feature-based (ability to deal with outlying matches for **robustness** using e.g. the RANSAC procedure, the Hough transform, robust estimators, etc.) methods. In addition, the task of matching blocks naturally lends itself to parallel implementation [8].

However, in this classical setting, point matching is based on the comparison of blocks whose relative positions in the two images are identical up to a **3D translation** defined on the **discrete** underlying image grid. These methods are thus likely to fail when dealing with large rotations, and to be unable to capture subvoxel displacements. In this paper, we propose to tackle these two problems by replacing the local, exhaustive search of optimal translations by a local, numerical search of optimal rigid-body transformations. Note that within this setting, by nature, the blocks in one of the two images are no longer defined on a discrete grid. The global optimal rigid-body transformation is then defined using least trimmed squares (LTS) regression from the locally estimated rigid-body transformations. We describe this new block-matching registration algorithm in Section 2. We then evaluate its accuracy, robustness, capture range and run time on real and synthetic data in Section 3, demonstrating its superiority over the classical block-matching algorithm. We conclude and give some perspectives in Section 4.

2. METHODS

2.1. The classical block-matching algorithm

The basic idea underlying the classical block-matching rigid-body registration algorithm is to iterate between the following

two steps, after the *floating* image has been resampled to share the same image grid as the *reference* image:

1. Matching homologous points in the two images by maximizing a similarity measure between intensities of subimages (hereafter termed blocks) around these points; the blocks to be compared are related via a discrete translation on the image grid, and located in a given neighborhood (hereafter termed search window).
2. Computation of the global transformation best superposing the points in the floating image with their counterparts in the reference image, and application of this transformation to the floating image.

The algorithm stops when two subsequently estimated global transformations are sufficiently close to each other. This algorithm as such is very generic and modular, and further choices must be made for each of these modules:

Similarity measure. The similarity measure can be designed to cope with monomodal (e.g. sum of squared differences, correlation coefficient) or multimodal (e.g. mutual information, correlation ratio) problems [9].

Blocks. The block-matching step relies on a set of four (vector) parameters, given the 3 directions x, y, z: the size of the blocks in both images ($N = (N_x, N_y, N_z)$), the grid step size for the blocks in the reference ($\Delta = (\Delta_x, \Delta_y, \Delta_z)$) and the floating ($\Sigma = (\Sigma_x, \Sigma_y, \Sigma_z)$) image, and the size of the search window in the reference image ($\Omega = (\Omega_x, \Omega_y, \Omega_z)$). It is common to choose these four parameters within a l-level multiresolution scheme. The blocks must be large enough for the similarity measure to be meaningful.

Optimization algorithm to match the blocks. A full (exhaustive) search is the common approach for this purpose. Other strategies have been extensively investigated for video compression [10], but have not been widely applied in medical image processing.

Optimization algorithm to compute the global transformation. Once point-to-point correspondences have been established, the global transformation is estimated as that minimizing the sum of squared distances between them, which can be solved in closed-form using one out of several well-known techniques [11]. Robust alternatives to the least-squares method can be used such as M-estimation or LTS regression [12].

2.2. Some problems, and a solution

This simple iterative algorithm is reminiscent of the popular ICP algorithm to register surfaces [13]. It suffers from the same flaw: the registration is likely to fail when the displacement between the two images is large, and especially when a large rotation is involved. In this case, the assumption that simple local translations are sufficient to recover the global rigid-body transformation turns out to be inadequate. Moreover, the discrete nature of the search procedure is not compatible with the continuous nature of the problem, and such an algorithm is unlikely to recover subvoxel displacements (e.g. one can think of a simple translation of half a voxel in one of the three directions x, y or z).

Our idea is to modify the block-matching strategy (first step of the algorithm) by replacing the search over local, discrete, grid-based translations by a search over local, rigid-body transformations. Then we propose to estimate the optimal global rigid-body transformation directly from those local rigid-body transformations. With these two modifications, we aim to recover large displacements, and especially some involving large rotations, with a subvoxel accuracy.

As a consequence of this new strategy, blocks in the floating image are no longer constrained to lie on the image grid, and the voxel values within these translated and rotated blocks must be interpolated each time the similarity measure is estimated, as opposed to within the simple translation setting. This results in a computationally expensive scheme, and we propose to use an efficient numerical optimization technique to perform this block-matching. We previously showed the NEWUOA algorithm to outperform classical algorithms (Powell's and Nelder-Mead's algorithms) in terms of speed, with similar robustness and accuracy for global rigid-body registration of MR volumes [14]. This is why we choose it here. In essence, NEWUOA is based on successive approximations of the similarity measure by quadratic functions, whose maxima can be computed analytically [15].

Once the optimization has been performed, we utilize the local rigid-body transformations R_i, $i = 1 \dots n$, to compute a global rigid-body transformation. If we define the residual r_i as $r_i \equiv ||\log(R) - \log(R_i)||$, where $||.||$ is the Frobenius norm, then the optimal transformation $\tilde{R}$ is estimated as $\log \tilde{R} = \arg\min_R \sum_{i=1}^{n-h} (r^2)_{i:n}$, where $0 \leq h \leq [n/2]$ and $(r^2)_{1:n} \leq \ldots \leq (r^2)_{n:n}$ are the ordered squared residuals. $\tilde{R}$ can be found iteratively [16], and each iteration m is a simple least-squares problem with a closed-formed solution: $\tilde{R}^m = \exp[(\sum_{i=1}^{n-h} \log R_{i:n}^{m-1})/(n-h)]$, where $R_{i:n}^{m-1}$, $i = 1 \dots n$, are the local rigid-body transformations ordered according to their squared residuals at iteration $m-1$ [17].

In the following, we compare this new strategy with the classical one (e.g. using full search over translations and LTS regression on the matched points) and with a third, hybrid algorithm where we use the NEWUOA algorithm to compute the translations, and LTS regression on the matched points. These three algorithms use the correlation coefficient (CC) as the similarity measure, in both mono- and multimodal contexts. We prefer it over the mutual information (MI) for two reasons. First, the MI requires a large number of samples to be meaningful. Second, the hypothesis of a linear relationship between intensities is valid when there are at most two tissues in the blocks, which is the case for most of the blocks when they are small enough; those where this assumption does not hold are likely to be eliminated from the estimation of the global transformation owing to the LTS regression.

3. VALIDATION & RESULTS

We evaluate the three block-matching strategies in terms of accuracy, robustness, capture range and run time, on three datasets composed of two images I_1 and I_2 each. Random rigid-body displacements are simulated for Datasets 1-2, while the ground truth transformation is known for Dataset 3.

Implementation details. The three evaluated block-matching algorithms use: exhaustive search of translations - FS, NEWUOA search of translations - N-Tr, and NEWUOA

search of rigid-body transformations - N-Rig. For all these algorithms, we use the same set of parameters for our experiments: $N = 7$ voxels, $\Delta = 5$ voxels, $\Omega = 2$ voxels, $\Sigma = 1$ voxel, $l = 3$ levels of pyramid and $h = 70$ %. NEWUOA also requires specific parameters: it is not using Ω and Σ but rather a search radius for the transformation parameters. We choose this search radius to be similar to exhaustive search parameters: 2 voxels for translation and 5 degrees for rotation (when applicable). The resampling of the floating image to compute the CC is done using trilinear interpolation.

Random experiments. For Datasets 1-2, I_1 and I_2 are initially aligned; we generate 150 random rigid-body transformations T with translations in the range $[0, 10]$ mm and rotations in the range $[0, 45]$ deg. around each of the three image axes, we apply $T^{1/2}$ to the first image I_1, and $T^{-1/2}$ to the second image I_2; this gives two misaligned images defined as, respectively, $\tilde{I}_1(v) = I_1(T^{-1/2}(v))$ and $\tilde{I}_2(v) = I_2(T^{1/2}(v))$, where v are the voxel positions. The resampling uses trilinear interpolation. We also add white Gaussian noise (std: 10% of the average image signal) to both $\tilde{I}_1$ and $\tilde{I}_2$.

Warping indices. We define the initial warping index $\omega_i = \sum_v ||T(v) - v||/\#(I_1)$; ω_i measures the misalignment of $\tilde{I}_1$ (reference image) and $\tilde{I}_2$ (floating image). After registration of $\tilde{I}_2$ onto $\tilde{I}_1$, giving the transformation F, we define the final warping index $\omega_f = \sum_v ||T(v) - F(v)||/\#(I_1)$; ω_f measures the registration error.

Evaluation metrics. The registration is considered a success if ω_f is lower than the voxel size. We define the **capture range** of the algorithm as the maximum ω_i for which it succeeds. We define its **accuracy** as the average ω_f for which it succeeds. We define its **robustness** as the percentage of random experiments being a success. We also report its average **run time** (on one core of a Xeon 3.0 GHz PC).

Dataset 1. The first dataset utilized in our experiments is from the BrainWeb MRI simulator[1]. We consider the T1-weighted vs T2-weighted MRI (size: 181×217×181, voxel size 1 mm^3) with no noise (as it is added in our experiments).

Dataset 2. The second dataset is a real dataset coming from an healthy subject. It is a dual echo T2/PD sequence (size: 192×256×44, voxel size 1×1×3 mm^3). As these two images are acquired simultaneously, they are natively aligned.

Dataset 3. The last dataset is a real clinical dataset coming from the R.I.R.E initiative for the comparison of multimodal registration algorithms[2]. It consists of an MR image (size: 256×256×26, voxel size 1.25×1.25×4 mm^3) and of a PET image (size: 128×128×16, voxel size 2.59×2.59×8 mm^3). These two images are acquired with fiducial markers providing objective landmark points (available from the website) from which the true transformation is computed.

Results. Our results on Dataset 1 are summarized in Table 1. This table demonstrates that strategies based on translation search (FS and N-Tr) have similar difficulties to deal with large transformations and achieve a relatively poor robustness. On the contrary, searching for local rigid-body transformations allows for a large capture range associated with a better robustness. In addition, utilizing NEWUOA for the optimization leads to fast computation times, making N-Rig as fast as FS even though local rigid-body transformations are optimized rather than simple discrete translations.

	Cap. range	Acc.	Rob.	Time
FS	84.0	0.14 ± 0.04	58.7	566
N-Tr	74.2	0.10 ± 0.03	44.7	270
N-Rig	101.6	0.59 ± 0.13	84.0	582

Table 1. Experiments on BrainWeb Simulated Dataset. Reports of capture range (mm), accuracy (mm), robustness (%) and average computation time (seconds) for the 150 simulated transformations on Dataset 1.

In addition, we report results for Dataset 2 in Table 2. These results confirm the results obtained on Dataset 1. The use of NEWUOA to optimize rigid-body transformations for each block leads to the best results allowing for a better accuracy than with exhaustive search block-matching, a larger capture range and a better robustness than with both exhaustive search (FS) and NEWUOA translation search (N-Tr). The proposed method is therefore able, thanks to the use of local rigid-body transformations, to recover larger transformations compared with other block-matching techniques.

	Cap. range	Acc.	Rob.	Time
FS	97.6	0.51 ± 0.08	64.7	663
N-Tr	103.4	0.12 ± 0.16	82.7	112
N-Rig	119.0	0.57 ± 0.10	99.3	208

Table 2. Experiments on T2/PD Real Dataset. Reports of capture range (mm), accuracy (mm), robustness (%) and average computation time (seconds) for the 150 simulated transformations on Dataset 2.

This is further illustrated in Fig. 1 on a difficult registration case with a large rotation. In this figure, we can see that only the combination of NEWUOA for optimization and the search for local rigid-body transforms is capable of recovering the transformation accurately.

Finally, we present the results obtained on Dataset 3 in Fig. 2, only for FS and N-Rig for the sake of concision. This figure illustrates that N-Rig is able to recover the transformation accurately between the two images whereas FS is not as accurate (difference visible close to the intersection of the green and red axes on Fig. 2). This is confirmed by the quantitative results computed with respect to the true transformation. The initial ω_i is of 27.9 mm. FS achieves a warping index ω_f of 9.8 mm, while N-Rig achieves 3.6 mm which is below the largest voxel size of the MR image. N-Rig is therefore able to achieve sub-voxel accuracy for this dataset while classical FS is not.

4. DISCUSSION & CONCLUSION

We have presented a new block-matching strategy for robust rigid-body registration of multimodal images. It is based on

[1] http://www.bic.mni.mcgill.ca/brainweb

[2] http://www.insight-journal.org/rire

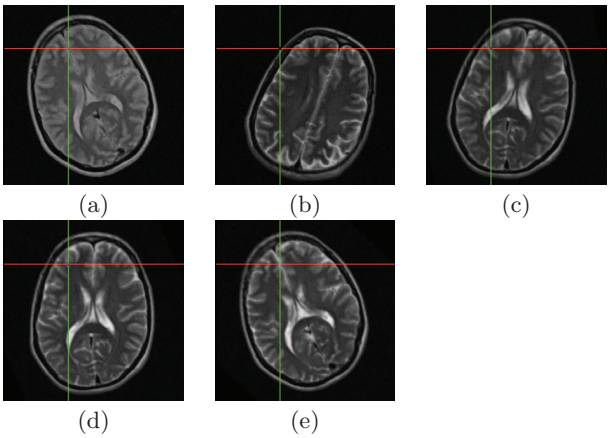


Fig. 1. **Registration Results of Real T2/PD Images**. Illustration of results on a difficult case of T2 / PD simulated transformation. (a): reference image, (b): floating image, (c,d,e): registered using respectively FS, N-Tr and N-Rig.

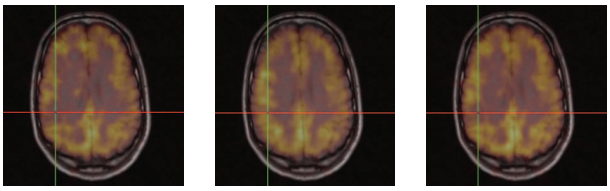


Fig. 2. **Registration Results of Real MR / PET Images**. PET image superimposed on the MR image after resampling with (a): the true transformation, transformations estimated with (b): FS, (c): N-Rig.

the optimization of local rigid-body transformations for each block thanks to the NEWUOA algorithm. We have demonstrated on simulated and real datasets that this approach allows one to overcome the drawbacks of the classical exhaustive search block-matching: correspondences are found in a non-discrete fashion, and rotations may be better recovered. In addition, searching for local rigid-body transformations allows the algorithm to converge faster, which combined with the NEWUOA optimization leads to faster computation times for N-Rig than for FS.

We plan to investigate other optimization algorithms such as Powell's bounded optimization algorithm BOBYQA [18] as it could allow one to constrain block displacements in a meaningful range. Finally, this algorithm may be easily extended to find affine transformations.

5. REFERENCES

[1] J.B.A. Maintz and M.A. Viergever, "A survey of medical image registration," *MedIA*, vol. 2, no. 1, pp. 1–36, 1998.

[2] B. Zitová and J. Flusser, "Image registration methods: a survey," *IVC*, vol. 21, no. 11, pp. 977–1000, 2003.

[3] C.A. Pelizzari, G.T. Chen, D.R. Spelbring, R.R. Weichselbaum, and C.T. Chen, "Accurate 3-D Registration Of CT, PET, And Or MR Images Of The Brain," *J. Comp. Assist. Tomography*, vol. 13, no. 1, pp. 20–26, 1989.

[4] W.M. Wells, P. Viola, H. Atsumid, S. Nakajimae, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *MedIA*, vol. 1, no. 1, pp. 35–51, 1996.

[5] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE TMI*, vol. 16, no. 2, pp. 187–198, 1997.

[6] J. Jain and A. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. on Communications*, vol. 29, no. 12, pp. 1799–1808, 1981.

[7] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block Matching: A General Framework to Improve Robustness of Rigid Registration of Medical Images," in *MICCAI'2000*, 2000, pp. 557–566.

[8] S. Ourselin, R. Stefanescu, and X. Pennec, "Robust Registration of Multi-modal Images: Towards Real-Time Clinical Applications," in *MICCAI*, 2002, pp. 140–147.

[9] A. Roche, G. Malandain, N. Ayache, and S. Prima, "Towards a Better Comprehension of Similarity Measures Used in Medical Image Registration," in *MICCAI'1999*, 1999, pp. 555–566.

[10] S.I.A. Pandian, G.J. Bala, and B.A. George, "A study on block matching algorithms for motion estimation," *International Journal on Computer Science and Engineering*, vol. 3, no. 1, pp. 34–44, 2011.

[11] D.W. Eggert, A. Lorusso, and R.B. Fisher, "Estimating 3-D rigid body transformations: a comparison of four major algorithms," *Machine Vision and Applications*, vol. 9, no. 5-6, pp. 272–290, 1997.

[12] P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*, Wiley Series in Probability and Mathematical Statistics, 1987.

[13] P.J. Besl and N.D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, 1992.

[14] N. Wiest-Daesslé, P. Yger, S. Prima, and C. Barillot, "Evaluation of a new optimisation algorithm for rigid registration of MRI data," in *Medical Imaging 2007: Image Processing*, 2007.

[15] M.J.D. Powell, "The NEWUOA software for unconstrained optimization without derivatives," in *Large-Scale Nonlinear Optimization*, vol. 83 of *Nonconvex Optimization and Its Applications*, pp. 255–297, 2006.

[16] P.J. Rousseeuw and K. Driessen, "Computing LTS Regression for Large Data Sets," *Data Min. Knowl. Discov.*, vol. 12, no. 1, pp. 29–45, 2006.

[17] V. Arsigny, O. Commowick, N. Ayache, and X. Pennec, "A Fast and Log-Euclidean Polyaffine Framework for Locally Linear Registration," *JMIV*, vol. 33, no. 2, pp. 222–238, 2009.

[18] M.J.D. Powell, "The BOBYQA algorithm for bound constrained optimization without derivatives," Tech. Rep., University of Cambridge, UK, 2009.