MODULAR LINEAR ICONIC MATCHING USING HIGHER ORDER GRAPHS

Vivien Fecamp[†‡] *Aristeidis Sotiras*[⋆] *Nikos Paragios*[†‡]

[†] Ecole Centrale Paris, Centre de Vision Numrique, Chatenay Malabry, France
[‡] INRIA GALEN Team, Paris Saclay, France
[⋆] University of Pennsylvania, Section of Biomedical Image Analysis, Philadelphia, USA

ABSTRACT

We introduce a novel approach to tackle iconic linear mapping between two images. We adopt a grid-based parametrization of the deformation field that is encoded by a higher order graphical model. In the proposed formulation, latent variables correspond to local grid displacement vectors and unary potentials locally quantify the level of alignment between the two images. Higher order constraints that involve third and forth order potentials, enforce the linearity of the resulting transformation. The resulting formulation is modular with respect to the image metric used to evaluate the correctness of mapping as well as with respect to the nature of the linear transformation (rigid, similarity, or affine). Inference on this graph is performed through dual decomposition. Comparison with classic algorithms demonstrates the potential of our approach.

1. INTRODUCTION

Deformable medical image registration has been the main focus of research, leading to important methodological advances and novel algorithms [1]. Linear registration despite being an important step in any medical image processing pipeline, has not profited by the many recent technological advances in the field.

Current state of the art in linear registration includes local and global methods. Local methods establish a subset of image correspondences and then optimally fit the best linear model using inference techniques [2]. Global methods [3] use conventional similarity metrics like Mutual Information (MI) or Sum of Absolute Differences (SAD) and seek the parameter set that corresponds to their lowest potential that is often determined through a gradient-driven optimization method. Their main strength is computational efficiency, while their main limitations include lack of robustness (sensitive to the initial conditions), lack of modularity (optimization depends on the linear model) and the fact that they can handle only differentiable objective functions. Simplex methods [4], cutting planes methods [5] or more recently discrete optimization ones [6] overcome to some extend the aforementioned limitations while suffering from computational complexity and precision.

Despite the fact that discrete optimization spans a great range of applications, it is rarely used to tackle global linear registration. To the best of our knowledge, only [6] has attempted to estimate global transformations with the use of Markov Random Fields (MRF). This is due to the fact that a global approach leads to a very complex graph if one wants to link every node to the node(s) encoding the transformation parameters, as they did in [6]. Here, we use the MRF theory and define a set a local higher order constraints to ensure the global displacement of the image is a linear transformation. The search space is thus smaller and tractable. We propose different higher-order constraints for varying degrees of freedom of linear transformation models. Our main contribution lies in showing how local third-order constraints can be used to ensure global transformation properties. Dual decomposition is used to infer the optimal displacements.

The rest of the paper is organized as follows: in Sec. 2, the construction of the graphical model is presented, the topology of the graph, the energy as well as the different potentials. The higher-order inference technique is explained in Sec. 3. In Sec. 4, the experimental validation on real 3D data is presented, with comparison with state of the art algorithms. The paper concludes in Sec. 5 with a discussion of the proposed method and some suggestions for future work.

2. METHOD

Given a source image and a target image, we seek to estimate the linear transformation that best aligns the source image to the target. The transformation is parametrized through the use of a deformation grid superimposed onto the source image. The deformation of the source image is governed by the movement of the control points. The goal is to find the displacements of the control points so that the global transformation of the whole image is an affine transformation. To find the displacement of a control point, we compare a patch of the target image centered at this control points with patches of the source image around the corresponding control point.

We use higher order MRFs to formulate our problem in a discrete context. The problem is represented by a hypergraph

$G = (V, C)$, where V denotes the set of nodes that encode the latent variables (grid node displacement) and C the set of cliques (or hyperedges) that encode higher-order interactions between the variables. V forms a grid of control points superimposed to the source image we want to register. Let $L = \{l_1, \dots l_n\}$ be the set of labels, where each element corresponds to a potential displacement vector. The algorithm associates a label l_p to each control point p, so that the final result best aligns the two images. The energy of the MRF can be written as:

$$E_{MRF} = \sum_{p \in V} U_p(l_p) + \sum_{c \in C} H_c(l_c), \tag{1}$$

where l_p is the label corresponding to the node p, $U_p(l_p)$ denotes the *unary potentials*, $l_c = \{l_p, p \in c\}$, is the set of labels assigned to the nodes in the clique c and $H_c(l_c)$ denotes the *higher order potentials*.

2.1. Unary potentials

Let p be a control point and l_p be the label associated to the control point p. Let B_p be the patch corresponding to the center point p in the source image, and B_{l_p} the patch corresponding to the image point p translated by the displacement vector l_p, in the target image. The unary potential quantifies how well the patch B_p matches the patch B_{l_p}:

$$U_p(l_p) = \rho(B_p, B_{l_p}), \tag{2}$$

where ρ, is an arbitrary intensity-based similarity measure. Being discrete, the proposed model is gradient-free and can encompass a wide choice of similarity measures, from SAD to statistical measures for multimodal registration like MI [7].

2.2. Higher order potentials

In order to constrain the global linear property of the transformation through local constraints, we exploit the property of linear transformations to preserve the barycentre. Let us denote by (p, q, r) three aligned control points in a clique c and (l_p, l_q, l_r) their respective labels. The condition (P) can be written:

$$\vec{l_p} + \vec{l_r} - 2 * \vec{l_q} = \vec{0}. \tag{3}$$

In order to guarantee that the transformation is linear, the violation of the condition (P) must be penalized. Thus, the higher potential is defined as:

$$H_c(l_c) = \Phi(\vec{l_p} + \vec{l_r} - 2 * \vec{l_q}), \tag{4}$$

where $\Phi()$ is a cost function defined as:

$$\Phi(x) = \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{otherwise.} \end{cases} \tag{5}$$

The λ-clique potential encodes different kind of constraints depending on the number of degrees of freedom of the global linear transformation we seek to recover. Let us denote the 4 points of the λ-clique as s, t, u, and v (as shown in fig. 1) and $s' = s + l_s$, t', u', and v' their respective images. In all three cases (affine, rigid, and similarity), the condition (P) should hold true for the three points of the diagonal, s, t, and u. Additionaly, we have:

For similarity registration: The images of the points s, u, and v should form an isosceles right triangle, with the image of v being the vertex of the right angle,

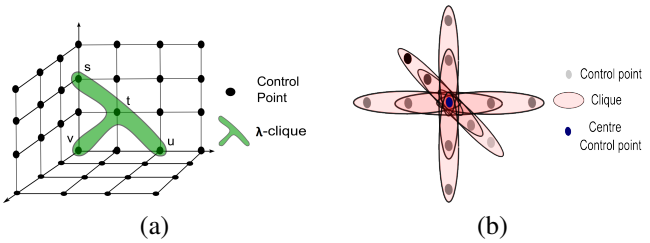
$$\begin{aligned} H_c(l_c) &= \Phi((\vec{s'} - \vec{v'}).(\vec{u'} - \vec{v'})) \\ &+ \Phi(\|(\vec{s'} - \vec{v'})\| - \|(\vec{u'} - \vec{v'})\|). \end{aligned} \tag{6}$$

The first part of the equation checks the orthogonality and the second part checks the two sides have the same norm.

For rigid registration: The isosceles triangle formed by the points s, u, and v should have the same size as the one formed by the respective images of the points. This leads to the definition of the following condition:

$$\begin{aligned} H_c(l_c) &= \Phi((\vec{s'} - \vec{v'}).(\vec{u'} - \vec{v'})) \\ &+ \Phi(\|(\vec{s'} - \vec{v'})\| - \|(\vec{u'} - \vec{v'})\|) \\ &+ \Phi(\|(\vec{s'} - \vec{v'})\| - \|(\vec{s} - \vec{v})\|). \end{aligned} \tag{7}$$

The additional third term penalizes differences in size between the triangle before and after the transformation.



(a) (b)

Fig. 1. The different cliques in 3D: (a) A λ-clique contains 4 control points forming a λ-shape. There is one λ-clique on each face of the cube. (b) The 9 cliques containing the central control point in blue.

2.3. Graph construction

Let us now detail the graph construction. The set of cliques C contains all the triplets consisting of collinear and neighbours points along each dimension of the grid. C also contains λ-cliques which are quadruplets, containing 3 points along a diagonal of the grid and the corner point, such that the 4 points form a T (see fig. 1). The presence of λ-cliques is necessary for the global properties of the linear transformation to hold. Due to lack of space, the proof is omitted.

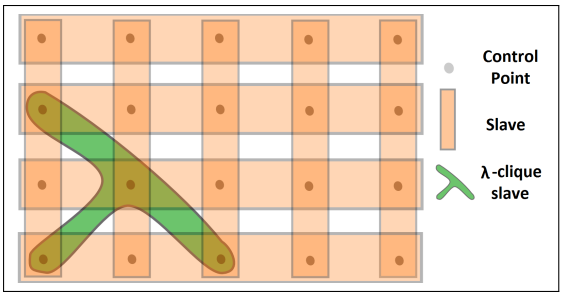


Fig. 2. The decomposition of the original problem in slave subproblems in 2D: one for each line, one for each λ-clique.

In 2D, each point belongs to three horizontal cliques and three vertical cliques. The graph contains one λ-clique positioned in one corner.

In 3D, each point belongs to three cliques along each axis (see fig. 1). The grid forms a cube and each face contains one λ-clique, therefore C contains 6 λ-cliques. Each λ-clique is similar to the one defined in the 2D case.

3. TECHNICAL DETAILS

3.1. Optimization method

To solve the MRF, we use Dual Decomposition [8]. Here, the graph is decomposed into trees (mostly chains) that constitute the set of subproblems or slaves. The only requirement for the choice of the subproblems is that they cover (at least once) every node and hyperedge of the hypergraph G. In our case, one slave problem is defined for each line parallel to a coordinate axis, and one slave for each λ-clique. An example of the different slaves in 2D is illustrated in fig. 2. One slave is a tree, while the rest are chains. Thus they can be solved easily and efficiently due to their simplicity.

3.2. Implementation details

We adopt a multiresolution approach, where the registration problem is solved in a coarse-to-fine manner. At each resolution level, we perform multiple iterations by successively refining the label space. The deformation grid is reset after each iteration and the resulting displacement fields are incrementally composed on the image.

The grid size contains 3^d points, where d is the image dimensionality ($d = 3$ for 3D images here), and is increased to 5^d points for the fine level of resolution. The label set is composed of 5^d displacement vectors. These displacement vectors are regularly spaced on a grid centered at the 0-displacement vector. The maximal length of the displacement vectors is 0.4 multiplied by the distance between two control points along each axis. This length is iteratively reduced, such that first iterations account for large displacements while later iterations account for small displacements. We used up to 8 iterations in our experiments.

Our algorithm is implemented in C++. The results were obtained using a 64 bits machine with a Intel Xeon W3670 processor and 16 GB of RAM.

Image	DICE	Sensitivity	Specificity
Before registration	70	71	95
Our registration	81	84	96
MedInria registration	71	73	95

Table 1. Results of the affine registration evaluated on manual segmentations of calf muscles MRI. Given is the DICE score, the sensitivity and the specificity.

4. EXPERIMENTAL VALIDATION

4.1. Comparison with MedInria

First, we used an annotated 3D MR-T1 (Magnetic Resonance images produced by spin-lattice relaxation) calf muscle dataset in order to compare the proposed method against the publicly available image analysis MedInria software. We report the DICE, the sensitivity and the specificity of the segmentations, before and after registration. We used the algorithm named Baloo [9], for affine registration, with the SSD similarity measure, with 5 iterations, 1 "coarsest pyramid level" and 1 "finest level".

The results are shown in table 1 and fig. 3 compares the repartition of the DICE for the two methods. Those results show that our algorithm performs better than MedInria registration in this dataset. The execution time was about 3 seconds for MedInria, and about 8 seconds for our algorithm.

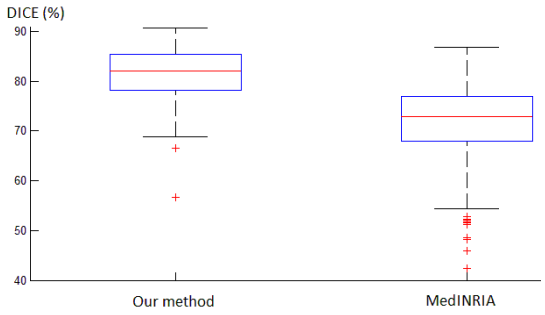


Fig. 3. Comparison of the DICE between our method and MedInria software. The proposed algorithm is more accurate than MedInria.

4.2. RIRE database

We test our method performing a series of 3D multimodal registrations of brain images (CT (Computed Tomography),

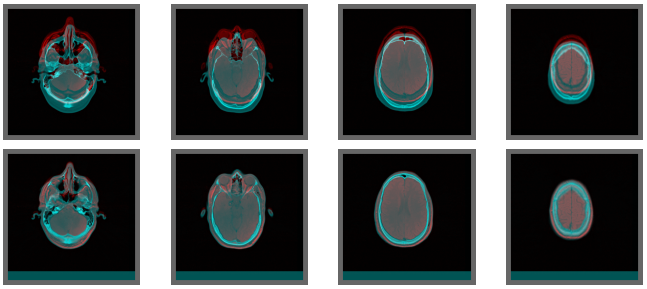


Fig. 4. Qualitative evaluation of registration. Source and target images are shown using different color bands. Axial views before (top) and after (bottom) registration are shown. The images have been successfully aligned.

MR-PD (Proton Density), MR-T1, and MR-T2 (spin-spin relaxation)) provided by *The Retrospective Image Registration Evaluation Project* (RIRE)[1]. The CT images have a resolution of $512 \times 512 \times 29$ and a physical voxel size of $0.65 \times 0.65 \times 4$ mm, while the MR images have a resolution of $256 \times 256 \times 26$ and a voxel size of about $1.25 \times 1.25 \times 4$ mm. We performed CT to MR registrations for 7 patients (Patient001 to Patient007, with a total of 21 registrations). We employed a histogram-based estimation of Normalized Mutual Information (NMI) using 32 bins as the similarity measure. We show in fig. 4.2 a slice of the 3D volumes before and after registration. We also show in table 3 different results obtained with our algorithm for different multimodal similarity measures.

We compare our results with FLIRT (results provided by [10]), Elastix [11], Simplex [4], and another MRF method [6] in table 2. This last paper reports the results and implementation details for Simplex, Elastix and Zikic's method. The accuracy of the registration is evaluated by computing the Fiducial Registration Error (FRE) for 10 fiducial points. All results are available on the RIRE website. The running time for the proposed approach, when using NMI, was approximately 2 minutes. The other methods run on our machine in approximately 5 minutes. Let us note that we report here the previously published results for the other methods, and not the results that we obtained when running them. This is because we were not able to accurately reproduce them.

In their article, Zikic et al. [6] left out one image (P03 with MR-PD) because none of the methods could register it. We include it back for comparison. Our method is competitive with state of the art methods in terms of quality of results and speed. Our score is poor when registering the MR-T1 images because we significantly fail to align one image, which considerably increases the mean error.

[1] Available at: http://www.insight-journal.org/rire/

	Mean error				
Modality	Simplex	Elastix	Zikic	FLIRT	Ours
PD	3.889	4.174	3.901	2.267	2.180
T1	1.124	1.186	1.157	1.656	4.255
T2	1.931	1.903	1.724	2.227	1.923

Table 2. Results of the 3D multimodal CT-MR registrations on patient data of the RIRE database. We report the mean FRE (in mm) after registration using the Simplex, Elastix, Zikic MRF, FLIRT, and our method.

	Mean error		
Modality	NMI(32 bins)	NMI(16 bins)	MI(32 bins)
PD	2.180	2.258	2.260
T1	4.255	4.030	2.247
T2	1.806	4.419	2.250

Table 3. Results obtained with the proposed method for different metrics, and different number of histograms bins used for the metric estimation. The adopted similarity measure, NMI with 32 bins, is indeed the most suitable measure to our problem.

5. CONCLUSION

In this paper, we have presented a discrete MRF formulation to solve the problem of linear registration, using a local higher order graphical model with hidden variables being the displacement vectors of the control points. The proposed formulation is modular with respect to the similarity criterion and the nature of the linear transformation (rigid, similarity, affine and could be extended to projective). Moreover, it is computationally efficient by means of its relative local nature and the designed search space. The performance of the method on 3D multimodal medical data along with comparisons with state of the art methods demonstrate the interest of investigating such a research direction.

Acknowledgments This research was supported by a DGA-MRIS scholarship.

6. REFERENCES

[1] Aristeidis Sotiras, Christos Davatzikos, and Nikos Paragios, "Deformable Medical Image Registration: A Survey," Tech. Rep. RR-7919, Sep 2012.

[2] S Ourselin, A Roche, G Subsol, X Pennec, and N Ayache, "Reconstructing a 3D structure from serial histological sections," *Image and Vision Computing*, vol. 19, no. 1-2, pp. 25–31, 2001.

[3] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith, "Improved Optimization for the Robust

and Accurate Linear Registration and Motion Correction of Brain Images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, oct 2002.

[4] R. Shekhar and V. Zagrodsky, "Mutual information-based rigid and nonrigid registration of ultrasound volumes," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 1, pp. 9 –22, jan. 2002.

[5] E. Kokiopoulou and P. Frossard, "Minimum distance between pattern transformation manifolds: Algorithm and applications," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 7, pp. 1225 –1238, july 2009.

[6] Darko Zikic, Ben Glocker, Oliver Kutter, Martin Groher, Nikos Komodakis, Ali Kamen, Nikos Paragios, and Nassir Navab, "Linear intensity-based image registration by Markov random fields and discrete optimization.," *Medical Image Analysis*, vol. 14, no. 4, pp. 550–562, 2010.

[7] Paul Viola and William M. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[8] Nikos Komodakis, Nikos Paragios, and Georgios Tziritas, "MRF energy minimization and beyond via dual decomposition.," *PAMI*, vol. 33, no. 3, pp. 531–52, Mar. 2011.

[9] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," in *Third International Conference on Medical Robotics, Imaging And Computer Assisted Surgery (MICCAI 2000)*, A.M. DiGioia and S. Delp, Eds., Pittsburgh, Penn, USA, octobre 11-14 2000, vol. 1935 of *Lectures Notes in Computer Science*, pp. 557–566, Springer.

[10] Marc Modat, David M Cash, Pankaj Daga, Gawin P Winston, John S Duncan, and Sébastien Ourselin, "A symmetric block-matching framework for global registration," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2014, pp. 90341D–90341D.

[11] Stefan Klein, Marius Staring, Keelin Murphy, Max a Viergever, and Josien P W Pluim, "Elastix: a Toolbox for Intensity-Based Medical Image Registration.," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 196–205, Jan. 2010.