

AN INTERACTIVE INTENSITY- AND FEATURE-BASED NON-RIGID REGISTRATION FRAMEWORK FOR 3D MEDICAL IMAGES

A. Azar^{1,2}, C. Xu², X. Pennec¹, N. Ayache¹

¹Epidaure, INRIA, Sophia-Antipolis, France
{Xavier.Pennec, Nicholas.Ayache} @sophia.inria.fr

²Siemens Corporate Research, Princeton, USA
{Antoine.Azar.ext, Chenyang.Xu} @siemens.com

ABSTRACT

This paper proposes a new interactive hybrid non-rigid registration framework that combines any intensity-based algorithm with a feature-based component, using an iterative dual energy minimization. The resulting transformation combines both intensity-based and feature-based deformation fields. The feature matching exploits user-placed landmark pairs, and based on saliency and similarity measures, optimizes the correspondences in the neighborhood of each landmark. A dense feature-based deformation field is then generated using a thin-plate spline interpolation. Additionally, the framework allows user interactivity for live guidance of the algorithm in case of errors or inaccuracies.

We present three experimental results of our hybrid approach on lung, pelvis and brain datasets, and show that in each case, the registration benefited from the hybrid approach as opposed to its intensity component alone.

1. INTRODUCTION

The wide variety of clinical applications of image registration has led to several different kinds of algorithms. Among them, we can identify two categories: intensity-based (iconic) and feature-based (geometric). As these algorithms use different information and present different advantages and drawbacks, we present a hybrid framework combining the strengths of both iconic and geometric approaches [1, 2, 3]. In addition, we allow a human expert to use his knowledge to guide the registration in difficult cases, and thus bring interactivity to ensure correct results. This leads to a simple interface allowing a physician to specify corresponding features on two different images, and a framework that integrates this geometric information into a given iconic algorithm. As human experts are very efficient at defining corresponding regions but generally lack enough precision to specify exactly two corresponding voxels, we propose to extend the inputted features to regions and perform a local region registration at each iteration to increase the accuracy of correspondences. We present qualitative and quantitative results of our framework for monomodal registration of lungs, pelvis and brain.

2. HYBRID REGISTRATION FRAMEWORK

In [1], in the context of inter-subject brain non-rigid registration, an energy equation is introduced for combining an intensity-based dense field with landmark-based correspondences located on the brain's sulci:

$$E(C_1, C_2, T) = E_{sim}(I, J, C_1) + \sigma \|C_1 - T\|^2 + \sigma\gamma \|C_2 - T\|^2 + \sigma\lambda E_{reg}(T) \quad (1)$$

where C_1 is a dense intensity-based field, C_2 are landmark-based correspondences, T is the final dense deformation, and γ is a trade-off coefficient between intensity matching and landmark matching. The minimization of this energy is done in 3 steps:

- Minimize w.r. to C_1 , which leads to a new dense field,
- Minimize w.r. to C_2 via ICP, which gives new matches between landmarks that most closely fit T ,
- Minimize w.r. to T to find an optimal and smooth T that fits C_1 and C_2 .

The estimated transformation [1, 4] is a weighted average of the fitting of intensity and feature points:

$$T(x) = \lambda K * C_1(x) + (1 - \lambda) \frac{\sum_i \alpha_i [C_2(x_i) - T(x_i)] K(x - x_i)^2}{\sum_i K(x - x_i)} \quad (2)$$

The minimization of the energy defined in equation (1) and the resulting transformation field in equation (2) assume a large number of landmarks (typically 2000 points on the brain's sulcus), with unknown correspondences between two sets of landmarks. We modify this framework to suit the context of a sparse set of user-placed landmarks and user-defined correspondences. The ICP computation between the two sets of landmarks is now unnecessary, as we can place high confidence in the correspondences. However, we do not expect a user to be voxel-level precise while selecting a landmark, and thus propose to minimize a region similarity measure. We also extend the definition of C_2 from a set of matching features to a dense interpolated

deformation field. It is possible to incorporate this definition modification in the framework described in [1]; However, we are interested in this work in using any intensity-based algorithm with no modification, and improve it using an expert’s knowledge. We thus cannot use directly equation (1) as it ties T to each energy minimization. In order to compute C_1 and C_2 fields at each iteration independently of T , we propose an iterative dual energy minimization formulation.

- Minimize with respect to C_1 and C_2 :
 $E(C_1, C_2) = E_{sim}(I, J, C_1, T_N) + \sigma E_{feature}(S_I, S_J, C_2, T_N)$
- Minimize with respect to T_{N+1} :
 $E(T_{N+1}) = \sigma|C_1 - T_{N+1}|^2 + \sigma\gamma|C_2 - T_{N+1}|^2 + E_{Reg}(T_{N+1})$

With T_N being the transformation T at iteration N , T_0 being an initial null displacement field, S_I and S_J defined respectively as the set of landmarks on images I and J .

One practical consideration of using a dual energy minimization and removing the constraints on C_1 and C_2 is that it may lead to larger displacements and less stability in converging towards a solution compared to the single energy formulation; However, in this formulation, C_1 and C_2 are still determined by considering the transformation T at the previous iteration. As a result, their computation is straightforward, and if the time step is small enough between iterations, we believe this system will converge to the same minimum as a single energy minimization. The stability of the proposed minimization scheme is observed in our preliminary experiments.

The major advantage of this approach is being able to incorporate any registration algorithm in our framework with no modification to its implementation. The framework proposed in [1] does not allow this.

Modifying the approach in [4] for our application yields the following transformation, which is a weighted-average of the fitting of intensity and feature points:

$$T(x) = (1 - \lambda(x))[G_1 * C_1(x)] + \lambda(x)[G_2 * C_2(x)] \quad (3)$$

with G_1 and G_2 being regularization kernels for each deformation field. We define $\lambda(x) \in [0,1]$ a confidence map across the image defining the trust placed in the intensity-based deformation versus the feature-based field. We build our confidence map by placing 3D normalized Gaussians around each feature, with the Gaussian variance defined as the distance between a feature and its match:

$$\lambda(x) = \frac{\alpha \sum_i \sigma_i K_i(x - p_i)}{1 + \left[H\left(\alpha \sum_i \sigma_i K_i(x - p_i) - \frac{\alpha}{\sqrt{2\pi}} \right) \cdot \left(\alpha \sum_i \sigma_i K_i(x - p_i) - 1 \right) \right]} \quad (4)$$

With K_i being a Gaussian of variance $\sigma_i^2 = (p_i - q_i)^2$, α a “dimming” parameter that we initially set to $\alpha = \sqrt{2\pi}$, $H(x)$ the Heaviside step function, defined at $H(0) = 1$, and p_i and q_i respectively the source and destination landmarks. This confidence map is updated at each iteration as landmark positions evolve (see 2.1). Different confidence maps could be built based on prior knowledge or tensor imaging, and is reserved for future work.

A ping-pong effect is theoretically possible between iterations, as C_1 and C_2 could be pulling a same voxel in different directions. In our context of interactive registration though, it is easy in case of oscillations to increase or decrease on-the-fly the influence of a particular landmark, or to add or remove landmarks.

2.1. Region-based non-rigid registration

The first step in computing our C_2 field is finding optimal matches from the user-placed features. Manually placed landmarks have the advantage of insuring the quality of the correspondence and of benefiting from the knowledge of a human expert. Unfortunately, it remains difficult for a physician to precisely specify two corresponding points in two images. For this reason, we define spherical regions around each landmarks as in [5], and register one region with respect to the other. The use of regions allows iterative searches in the neighborhood of each pair of landmarks for optimal correspondences. These new landmark positions will drive our feature-based registration.

We propose here an algorithm based on saliency [6] (a local entropy-based information measure) and similarity measures to compute optimal correspondences between regions.

Define regions Ω_p around landmark p and Ω_q around landmark q
For every voxel in Ω_p :
compute saliency A and scale s
Label p' the voxel in Ω_p with the highest saliency
Define $\Omega_{p',s}$ as the spherical region of scale s around landmark p'
For every voxel in Ω_q :
label current voxel as q^*
define $\Omega_{q^*,s}$ as the spherical region of scale s around landmark q^*
compute similarity measure between $\Omega_{q^*,s}$ and $\Omega_{p',s}$
Label q' the voxel in Ω_q with the highest similarity measure

Table 1 – Algorithm to compute new feature positions around manually placed landmarks. Original landmarks are respectively noted by p and q in images I and J , and newly computed landmarks around p and q are noted by p' and q' .

Strong saliency in a region indicates high uniqueness: a salient region in image I can thus be associated with as few other regions as possible in image J (ideally only one, which defines the exact correspondence). A region with low

saliency bears more chances of being matched with a similar but incorrect region, such as background.

Saliency features are also well adapted to our landmark matching as it is possible to achieve scale, rotation and translation invariance. The translational part of the transformation is straightforward and is defined as the difference between feature centers. The scaling factor can also be computed as the ratio of the feature scales. In order to achieve rotation invariance, we perform as in [7] a local registration in order to estimate the best rotation parameters, using gradient descent. This rotation information can then be used to modify the deformation field generated around that landmark. It is left to future work to define a formal way to introduce this new field in our transformation.

The similarity measure used to compute the quality of a correspondence can be chosen depending on the type of images used (e.g. mono-modal or multi-modal). In our tests, we found the SSD gave the best results as all images are monomodal.

The algorithm presented in Table 1 is visually demonstrated in Figure 1. We manually placed landmarks in two pelvis images, and then rigidly registered the images. If the correspondences were correct, the landmarks would perfectly align. Note that this example has been run in 2D for presentation purposes, but the implemented algorithm performs in 3D, where human precision is much lower.

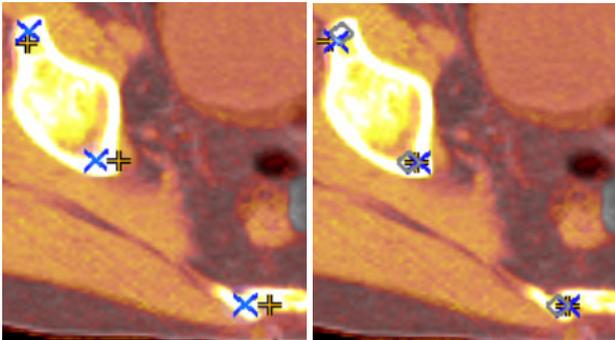


Figure 1 – Left: manually placed landmarks. Right: landmarks after one iteration of our algorithm. The + crosses represent the set of landmarks using saliency, and the x crosses represent the other set using similarity measures to match the + positions. The diamonds show the original placement of the x landmarks.

Figure 1 shows that after running one iteration of the algorithm, the landmarks are perfectly aligned.

Once the optimal matching step has been done, we perform a thin-plate spline interpolation [8] across the image. This yields our dense C_2 deformation field. The TPS has been chosen as it is very fast to compute with few landmarks (only one small matrix to invert), is guaranteed to displace a source landmark to its destination, and is always invertible.

2.2. Intensity-based non-rigid registration

It is possible to incorporate any intensity-based algorithm in our registration framework, and improve its results via user-input in the form of features.

In our research, we have successfully used two previously developed algorithms [9, 10], with no modification to their source code and very fast integration times. Indeed it is only necessary to input to the algorithm our two images, the current transformation, and collect as an output the C_1 deformation field. The results presented in the next section have been generated using an efficient intensity-based non-rigid registration algorithm [9].

3. RESULTS

Quantitative tests are provided by computing the SSD and SAD over the whole images. All images were initially rigidly registered. For the hybrid approach, only about 5-10 landmarks were used. For all cases, each iteration took less than a minute to compute (typically a few seconds). We denote in this section the intensity-based approach as IB.

3.1. Lungs dataset (128x128x128 voxels)

For this dataset, the SSD and SAD show only 1% improvement using the hybrid method. However, observing the images reveals that the IB algorithm was incorrectly thinning the tissue separating the lungs. As the phenomenon occurred, it was easily possible in the hybrid approach to avoid this thinning by inserting a few landmarks between iterations. Figure 2 shows results after 3 iterations for both approaches, with the landmarks used in the hybrid case.

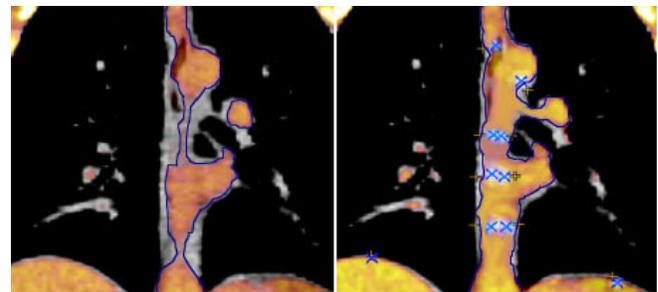


Figure 2 – Results for the lungs datasets. Left: IB, with obvious tissue thinning. Right: hybrid with correct tissue. Both datasets are fused, one in gray and the other in hot colors. We segmented the latter dataset for better visualization.

It is here apparent that allowing a physician some interactive input during a registration process can not only guide the algorithm towards a global minimum, but also provide important corrections for anatomical consistency.

3.2. Pelvis dataset (256x256x79 voxels)

Our quantitative tests for the pelvis datasets show equal results for both approaches. However, a visual examination shows again the IB approach failed to register correctly the bladder and fell in a local minimum, while this was avoided easily with a few landmarks in the hybrid case.

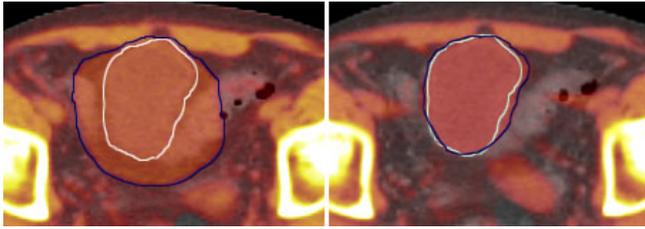


Figure 3 – Results for the pelvis dataset, with a close-up on the bladder (segmented). Left: IB. Right: hybrid, with a better superimposition of the bladder segmentations.

While a global energy may decrease, important anatomical areas in regions of interest might not be correctly registered at all in a pure IB approach. Our framework allows a physician to emphasize the registration in specific areas.

3.3. Brain dataset (256x256x100 voxels)

In this artificial test, we used a brain dataset, and applied to it a generated dense deformation field. This field was created by convoluting a sparse set of random deformation vectors (spanning 10% of the size of the image) by 3D Gaussians. Table 2 summarizes the results.

5 iter.	Initial	Intensity	Feature	Hybrid
SSD	109166	20461	70681	14202
SAD	165.4	68.2	123	60

Table 2 – Quantitative results on the brain datasets

It is clear the feature-only approach lacks the overall accuracy of an intensity-dependant algorithm. However, figure 4 reveals the IB algorithm, while accurate in most parts of the image, fell in a local minimum. The hybrid approach avoided this minimum and gave correct and accurate results with only a few landmarks to guide it.

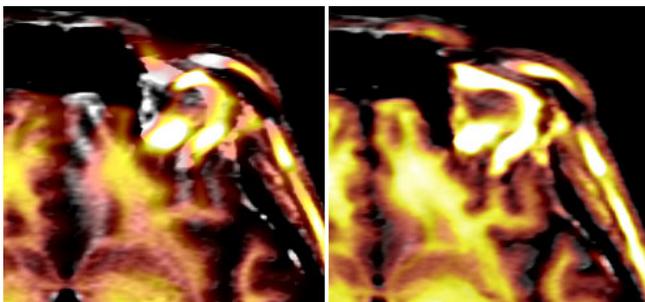


Figure 4 – Results for the brain dataset: Left: intensity-based after 8 iterations. Right: hybrid after only 5 iterations

4. DISCUSSION AND FUTURE WORK

While visual results show clear improvements in regions of interest, global quantitative results often fail to represent those benefits, and don't emphasize important localized corrections. We thus intend to develop a new local measure to better show the impact of user-based corrections. We should also note that improving the results of a powerful IB algorithm such as the one used here is a greater challenge to our hybrid approach than using a weaker algorithm. We plan to better formalize our energy minimization, and to incorporate in our TPS interpolation the rotation and scale components of the local region-based registration. We will also improve our confidence map using physiological and anatomical information. It is possible to extend our framework to incorporate any number of registration algorithms (as no implementation change is needed), and benefit from their own advantages. The resulting "meta-algorithm" should then yield better results than any algorithm alone and could be used for validation purposes. Finally, we intend to run more intensive tests using difficult datasets, such as multimodal or interpatient images. Indeed, the harder the registration, the more effective our hybrid approach becomes as human intelligence can guide it.

5. REFERENCES

- [1] P. Cachier et al., "Multisubject Non-rigid Registration of Brain MRI Using Intensity and Geometric Features", *MICCAI'01*, 734-742, October 2001.
- [2] P. Hellier, C. Barillot, "Coupling Dense and Landmark-based Approaches for Non Rigid Registration", *IEEE TMI*, 22(2): 217-227, 2003.
- [3] L. Collins, G. Le Goualher, and A. Evans, "Non linear cerebral registration with sulcal constraints", *MICCAI'98*, vol. 1496 in LNCS, 974-985. Springer, October 1998.
- [4] P. Cachier and N. Ayache, "Isotropic energies, filters and splines for vectorial regularization", *J. of Math. Imaging and Vision*, 20(3):251-265, May 2004
- [5] X. Huang et al., "Hybrid Image Registration based on Configurational Matching of Scale-Invariant Salient Region Features", *CVPRW'04*, Washington D.C., July 2004.
- [6] T. Kadir and M. Brady, "Saliency, scale and image description", *IJCV*, 45(2):83-105, November 2001.
- [7] D. Hahn, "A Practical Salient Region Based Hierarchical Method for Aligning 3D Medical Images", *Diploma Thesis*, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2005.
- [8] F. L. Bookstein, "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations", *PAMI*, 11(6):567-585, 1989.
- [9] C. Chef'd'hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multimodal image registration", *ISBI'02*, Washington, USA, 2002.
- [10] C. Guetter, C. Xu, F. Sauer and J. Hornegger, "Learning based non-rigid multi-modal image registration using Kullback-Leibler divergence", *MICCAI'05*, LNCS 3750, 255-262 .