

Hierarchical Multimodal Image Registration Based on Adaptive Local Mutual Information

Dante De Nigris¹, Laurence Mercier², Rolando Del Maestro³,
D. Louis Collins², and Tal Arbel¹

¹ McGill University, Centre for Intelligent Machines
{dante,tal}@cim.mcgill.ca

² McGill University, Dept. of Biomedical Engineering
{laurence,louis}@bic.mni.mcgill.ca

³ Montreal Neurological Institute and Hospital, McGill University
rolando.del_maestro@mcgill.ca

Abstract. We propose a new, adaptive local measure based on gradient orientation similarity for the purposes of multimodal image registration. We embed this metric into a hierarchical registration framework, where we show that registration robustness and accuracy can be improved by adapting both the similarity metric and the pixel selection strategy to the Gaussian blurring scale and to the modalities being registered. A computationally efficient estimation of gradient orientations is proposed based on patch-wise rigidity. We have applied our method to both rigid and non-rigid multimodal registration tasks with different modalities. Our approach outperforms mutual information (MI) and previously proposed local approximations of MI for multimodal (e.g. CT/MRI) brain image registration tasks. Furthermore, it shows significant improvements in terms of mTRE over standard methods in the highly challenging clinical context of registering pre-operative brain MRI to intra-operative US images.

Keywords: multimodal image registration, image-guided neurosurgery.

1 Introduction

Traditional multimodal image registration techniques rely on the assumption that features exposed in one modality (e.g. blobs, borders) will be similarly observable in a second modality, although possibly with a different intensity response. In this paper, we consider contexts where this assumption is fundamentally violated, as in image-guided neurosurgery, where the objective is to match pre-operative Magnetic Resonance Images (MRI) of patient brains to intra-operative ultrasound (US) images. Inherently, the two modalities convey very different information. US images indicate changes in acoustical impedance in the direction of the sound wave and therefore expose tissue boundaries and liquid-filled structures relatively well. They do not, however, provide distinctive intensity maps relating to different tissue types, as in MRI. This problem, immediately observable in Figure 1, poses serious challenges to most intensity based techniques, including traditional and

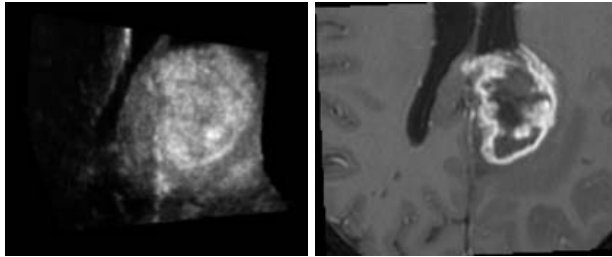


Fig. 1. US brain image acquired during neurosurgery and corresponding pre-operative MRI slice. Notice that the tumor appears as a noisy, diffuse bright blob in the US. In the MRI, the tumor has a bright boundary and a dark center.

normalized mutual information (MI) [1,2,3]. Previous work on MRI/US registration with intensity-based similarity metrics has focused on pre-processing the images to improve their intensity correspondences [4,5] and/or building a mask of the US image where an intensity-based similarity metric has a higher probability of finding correspondence[6].

In this work, we propose a registration framework which embeds a new localized similarity metric expressed as an orientation similarity measure based on a local approximation to MI. Other local MI metrics have been devised [7,8], however, our metric is adaptive, in that it can loosen or tighten its constraints depending on the modalities being registered and the scale at which the images are being registered. We also propose a computationally efficient method for estimating gradient orientations. Experiments with real multimodal (e.g. Computed Tomography (CT) and MRI) brain image registration tasks indicate that it outperforms both standard MI and other established orientation similarity metrics. Furthermore, it shows significant improvements over standard Normalized MI (NMI) in the highly challenging clinical context of registering pre-operative brain MRI to intra-operative US.

2 Method

2.1 Local Mutual Information - Previous Work

In [7], the authors developed a local similarity metric based on the analytical limit of MI as the window of observation approaches the voxel size. The expression for local MI (LMI) is obtained by modeling both images with a first-order Taylor expansion and is a monotonically decreasing function of the angle, θ , between the gradient orientations,

$$LMI(\theta) = C_d + \log_2 |\sin(\theta)| \quad (1)$$

where C_d is a constant that depends on the dimension of the image.

It is not possible to build an energy function by simply summing Eqn. 1 over a set of points in the image, since an extremum would appear whenever any of the points has a minimal inner angle. To circumvent this issue, the authors simplify the expression to,

$$LMI_2(\theta) = \frac{1}{2} \cos^2(\theta) \tag{2}$$

which has a smoother shape and whose energy function exhibits an extremum only when there is a collective coherence in terms of orientation similarity.

In related work [8], the authors preserve the dynamics of the original expression and eliminate the singularity by including an ϵ factor,

$$LMI_3(\theta) = \log_2(\epsilon + |\sin(\theta)|) \tag{3}$$

The two localized metrics have significantly different *coherence-selectivity* trade-offs. A coherent metric, such as Eqn. 2, can be loosely defined as one that is maximal when a majority of points show some degree of image correspondence. In contrast, a selective metric, such as Eqn. 3, has stricter correspondence constraints and therefore exhibits an extremum as soon as a few points comply with such constraints. A measure that favors coherence will tend to be smoother but less accurate because it will effectively average over all sampled pixels.

2.2 Adaptive Local Mutual Information

Our work builds on previous work [7,8] in which MI is locally approximated by a metric appropriately derived from the energy function. We propose an adaptive local orientation-based similar metric (ALMI) that addresses the coherence-selectivity trade-off directly as follows:

$$ALMI(\theta; K, \theta_c) = 2 - \frac{1}{1 + e^{-K(\theta - \theta_c)}} - \frac{1}{1 + e^{-K(\pi - \theta - \theta_c)}} \tag{4}$$

The selectivity of this smooth sigmoid-based function is adaptable by varying the curvature, K , and cutoff angle, θ_c . Figure 2 illustrates how the two previously proposed measures compare to various configurations of our proposed expression. An important advantage of ALMI is that it saturates smoothly to a maximum as $\theta \rightarrow \{0, \pi\}$. This property allows the energy function to reach a stable state as the optimizer approaches an extremum.

2.3 Adaptive Selectivity and Sampling

In addition to the metric proposed above, we propose an adaptive multiscale pixel selection scheme and a strategy to parametrize the similarity metric. Rather than adopting a standard low-pass image pyramid [9] in which the images are both blurred and down-sampled at each level, we use a smoothing image pyramid in which only the scale of Gaussian blurring is increased at each level. Figure 3 illustrates how the gradient orientations change with different blurring scales.

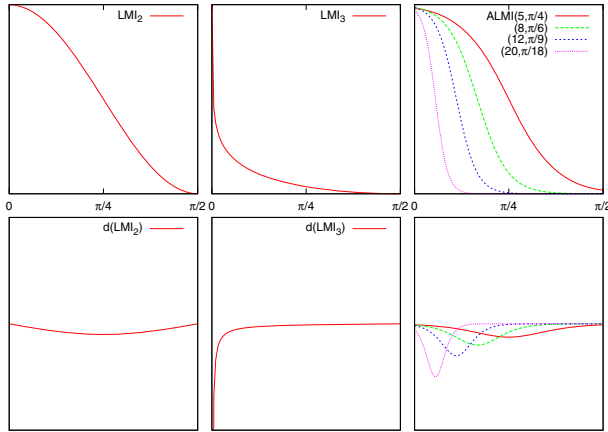


Fig. 2. The top row shows the two previous proposed localized measures and our proposed measure as a function of θ . The bottom row shows their respective derivatives.

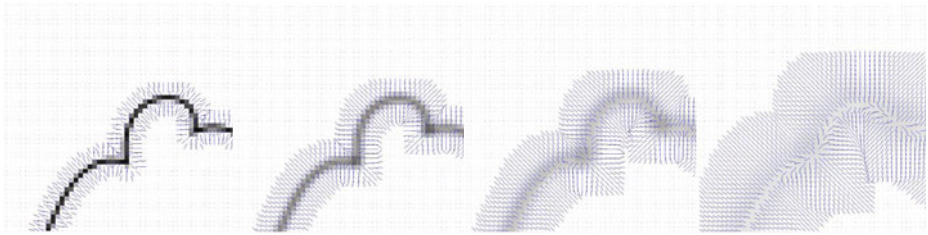


Fig. 3. Gradient Orientations (short blue lines) at different blurring scales, $\sigma \in 0, 1, 2, 4$

As the scale of diffusion is increased, the spatial support of a structure (i.e. the extent of image area where the gradient orientations accurately indicate the orientation of a structure of reference) tends to increase proportionately to the ratio of the current scale to the previous scale.

Further performance gains can be attained by focusing computations on a cleverly selected subset of voxels in the images. In this work, our conjecture is that the gradient magnitude is a valid indicator of the *reliability* of an estimated orientation and therefore we select high gradient voxels. The performance of our method is improved by adapting this sampling mask, as well as the parameters of the ALMI similarity metric in relation to the current Gaussian scale and the quality of the modalities being registered. We express the adaptive masking function $M(\mathbf{x})$ as,

$$M(\mathbf{x}) = \begin{cases} 1 & \text{if } |\nabla I_f(\mathbf{x})| > g(p_\sigma) \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

where $p_\sigma = \left(\frac{\sigma}{\sigma_1}\right)p_{\sigma_1}$ and $g(p)$ is the threshold value that captures the top p percentile of the fixed image gradient magnitude.¹

2.4 Cost Saving Approximations

A metric based on gradient orientations incurs a computational cost related to evaluating the gradient at each point of interest. We propose a simplifying implementation in which the gradients are computed only once for both the fixed and the moving image. Subsequent gradient orientations of the transformed moving image are estimated from the initial orientations.

In order to permit non-rigid registrations, we make the simplifying assumption that the transformation in the voxel neighborhood can be well approximated by a locally rigid transformation. Hence, we can estimate the gradient orientation of a point by multiplying the gradient orientation of the point prior to transformation with an estimated rotation matrix as follows:

$$\begin{aligned}\theta(\mathbf{x}) &= \angle(\nabla I_f(\mathbf{x}), \nabla I_m(T(\mathbf{x}))) \\ &\approx \angle(\nabla I_f(\mathbf{x}), R \cdot \nabla I_m(\mathbf{x}' = T(\mathbf{x})))\end{aligned}$$

Such a scheme eliminates the effect of intensity-interpolation artifacts and minimizes the expense of using high quality gradient operators.

3 Experiments

In this section, we present the results of evaluating our method, implemented within the Elastix toolbox [10], with two clinical datasets. Section 3.1 details the results of CT/MR rigid registration cases from a public dataset, while Section 3.2 details the results of rigid and non-rigid registrations with real MR and US images during image-guided neurosurgery.

3.1 Rigid Registration of CT/MR Brain Images

We evaluated the performance of our method with a series of CT/MR brain registration cases from the Retrospective Image Registration Evaluation (RIRE)² project. Since both modalities preserve structure and a large extent of that structure has low curvature, we chose a selective profile ($K = 20$, $\theta_c = \pi/18$) for all resolutions.

Table 3.1 lists the results of our proposed method with an adaptive masking techniques ($ALMI_m$) and without ($ALMI$). We also show the results of registration with previous proposed local metrics based on orientation similarity (LMI_2, LMI_3), a standard MI technique, a standard NMI technique and a

¹ In our experiments, we found that setting $p_{\sigma_1} = 0.1$ (i.e. sampling from the top 10% at the finest resolution) consistently yielded very good results.

² <http://www.insight-journal.org/rire>

more sophisticated technique which maximizes NMI with Generalized Partial Volume Interpolation (NMIGPV) [11].³ The mTRE is evaluated over the Volumes of Interest (VOI) of all cases. ALMI-based registrations on average show an improved accuracy over the previous proposed orientation similarity metrics. In addition, the metrics with high selectivity generally perform better than the coherent metric, LMI_2 . The standard MI and NMI methods exhibit relatively large mTREs, possibly because they do not embed any scheme to reduce the effect of interpolation-induced artifacts [12,3]

Table 1. Registration Results expressed in mTRE (in mm) of the VOIs of six patients

	MI	NMI	NMIGPV	LMI_2	LMI_3	ALMI	$ALMI_m$
	Intensity Based			Orientation Based			
CT/T1	3.20	3.43	0.87	2.31	1.42	1.14	1.00
CT/T2	1.90	5.03	1.11	2.87	1.56	1.45	1.39
CT/PD	4.11	3.88	0.90	2.36	0.82	1.07	0.91

3.2 MRI/US for Image Guided Neurosurgery

We apply our framework to the clinical context of image-guided neurosurgery where intra-operative US images are acquired during open craniotomies and matched to pre-operative MRI in order to correct for brain shift (linear and non-linear). Figure 1 shows extracted planes from a registered US and MR volume. Notice that apart from key structural components like the right lateral ventricle and the fissure, the ultrasound does not have a global intensity or orientation mapping with respect to the MR slice. However, we can argue that the US high gradient pixels do exhibit a loose orientation correspondence with the corresponding MR pixels.

Figure 4 illustrates our proposed similarity metric evaluated with four different configurations formed by a choice of a selective ALMI or a non-selective ALMI, and a choice of full sampling or sampling of the high gradient regions. The energy function between the US and MR slice of Figure 1 is evaluated in the parametric space of translational displacements in the x and y axis. We can observe that choosing a non-selective ALMI greatly reduces the number of local extrema. Furthermore, sampling only the high gradient regions of the US further reduces the number of local extrema, increasing the probabilities of a successful registration. Hence, we implement a strategy based on ALMI that relaxes the orientation similarity constraint as finer resolutions are approached.

Multimodal image registration experiments are performed on four tumor cases acquired at the Montreal Neurological Institute. The cases include: 1) high-grade glioma in the left frontal lobe, 2) high-grade glioma in the left occipital lobe,

³ The results of the rigid CT/MR registration were obtained directly from the RIRE website, where fellow researchers upload the results from their implementations using the same real datasets. We chose to compare our method against those results so as to avoid biases introduced by implementing other techniques.

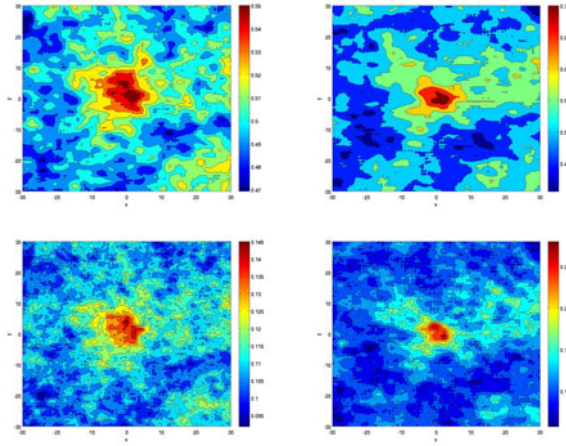


Fig. 4. The energy functions in the top row use a relaxed orientation matching constraint, evaluated with $(K = 5, \theta_c = \pi/4)$ while the energy functions on the bottom row use a strict orientation matching constraint, evaluated with $(K = 5, \theta_c = \pi/4)$. The energy functions in the left column evaluate all overlapping pixels, while the energy functions on the right column evaluate only the top 10% high gradient US pixels.

3) high-grade glioma in the right parietal lobe, and 4) low-grade glioma in the left temporal lobe. Each case includes a pre-operative gadolinium- enhanced T1-weighted MRI and a 3D ultrasound image acquired on the dura before starting the resection. Ultrasound images were acquired with an HDI 5000 (ATL/Philips, USA) using a P7-4 MHz phased array transducer at a depth setting of 6.5 or 8 cm. Each acquisition consisted in 200-600 tracked images that were reconstructed into a 3D volume at a resolution of $0.3 \times 0.3 \times 0.3$ mm. To validate registration, an expert (neurosurgeon) chose 7 to 11 corresponding anatomical points in both modalities. With every estimated transformation, we transform the locations of the anatomical points in the modality of the fixed image (US), evaluate the distance (i.e. error) from the corresponding anatomical points in the moving image (MR), and compute their mean.

Table 2 shows the results of the four cases after a rigid registration and after a subsequent non-rigid registration characterized by a B-Spline transformation. We compared our strategy against a standard multiresolution registration that maximizes NMI, and against a state-of-the-art method that pre-processes the US and MR volumes and then maximizes NMI [6]. Both NMI based registration techniques have relatively high registration errors in this domain. Only in cases 3 and 4, where there was only a small displacement, did the rigid registration show improvement. Non-rigid registrations led to consistent divergence from the ground truth. Conversely, our proposed approach manages to reduce the mTRE in all cases, three of them with an average accuracy below 2 mm. Furthermore, applying a non-rigid transformation further improved the accuracy in the first two cases, and slightly degraded the second two cases.

Table 2. MRI/US Registration Results expressed in mTRE (in mm) of expert selected anatomical points of interest

Case	Initial	NMI		Pre-processing + NMI		ALMI	
		Rigid	Nonrigid	Rigid	Nonrigid	Rigid	Nonrigid
1	10.50	14.75	17.12	16.03	13.81	1.87	1.40
2	10.86	17.52	20.80	13.94	13.94	2.96	2.53
3	2.62	1.98	5.70	2.31	4.28	1.63	1.86
4	3.24	2.91	7.44	3.66	8.07	1.93	2.06

4 Discussion

We have presented a new approach for multimodal image registration based on an adaptive localized similarity metric and an adaptive image masking scheme. We have provided basic guidelines for the choice of ALMI parameters in relation to the quality of the modalities being registered and the Gaussian blurring scale. Nevertheless, in an extended version of this work, we will seek to provide a more detailed analysis of the influence of different parameter selections. We also introduced a computationally efficient implementation that evaluates the image gradients only at the start of each resolutions optimization. We have shown improvements in accuracy and robustness across various modalities. We have also demonstrated that the approach is general enough to tackle a difficult problem such as MRI/US registration.

References

1. Maes, F., Collignon, A., Dirk, V., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging* 16, 187–198 (1997)
2. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. In: *IJCV*, pp. 16–23 (1995)
3. Pluim, J., Maintz, J., Viergever, M.: Mutual-information-based registration of medical images: a survey. *IEEE Trans. Med. Imaging* 22(8), 986–1004 (2003)
4. Arbel, T., Morandi, X., Comeau, R., Collins, D.L.: Automatic non-linear MRI-Ultrasound registration for the correction of intra-operative brain deformations. In: Niessen, W.J., Viergever, M.A. (eds.) *MICCAI 2001*. LNCS, vol. 2208, pp. 913–922. Springer, Heidelberg (2001)
5. Roche, A., Pennec, X., Malandain, G., Ayache, N.: Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information. *IEEE Trans. Med. Imaging* 20, 1038–1049 (2001)
6. Ji, S., Wu, Z., Hartov, A., Roberts, D.W., Paulsen, K.D.: Mutual-information-based image to patient re-registration using intraoperative ultrasound in image-guided neurosurgery. *Medical Physics* 35(10), 4612–4624 (2008)
7. Karaçali, B.: Information theoretic deformable registration using local image information. *IJCV* 72, 219–237 (2007)

8. Biesdorf, A., Wörz, S., Kaiser, H.J., Rohr, K.: Hybrid spline-based multimodal registration using local measures for joint entropy and mutual information. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5761, pp. 607–615. Springer, Heidelberg (2009)
9. Lindeberg, T.: Scale-space theory in computer vision. Springer, Heidelberg (1993)
10. Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J.: elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imaging* 29(1), 196–205 (2010)
11. Chen, H., Varshney, P.K.: Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation. *IEEE Trans. Medical Imaging* 22(9), 1111–1119 (2003)
12. Tsao, J.: Interpolation artifacts in multimodality image registration based on maximization of mutual information. *IEEE Trans. Med. Imaging* 22(7), 854–864 (2003)