GRAPH-THEORETIC IMAGE REGISTRATION USING PRIOR EXAMPLES

Mert Rory Sabuncu, and Peter J. Ramadge

Department of Electrical Engineering, Princeton University
Princeton, NJ 08544, USA
email: msabuncu,ramadge@princeton.edu

ABSTRACT

Information theoretic image and volume registration is currently of interest as a method for multi-modal alignment. It has been suggested that it is useful to incorporate information obtained from previous registrations into these methods to improve future registration performance. In this paper we examine how this can be done when using a graph theoretic estimator of entropy. Our main contribution is a method for incorporating prior information in a natural way and with minimal computational overhead into a registration measure based on a Euclidean minimal spanning tree estimate of entropy.

1. INTRODUCTION

Images of the same physical structure obtained through different sensing modalities, e.g. magnetic resonance (MR), computer tomography (CT), ultra-sound (US), etc, are often assumed to be well modelled through some unknown but fixed dependency of the image intensities. Since the images are taken using different sensors, in general they are not spatially aligned. Hence registration (spatial alignment) is performed to compare and/or fuse the images. Prior to registration, there are thus two forms of uncertainty: uncertainty in the modal relationship and uncertainty in the spatial alignment. Registration attempts to remove spatial uncertainty and in the process also reduces the uncertainty in the modal relationship.

Consider a situation where we need to register a sequence of multi-modal image pairs (I_1^k, I_2^k), $k = 0, 1, 2, \dots$. At time $k-1$ we have registered the image pairs (I_1^j, I_2^j) for $j = 0, \dots, k-1$. If the spatial alignment of two distinct pairs is independent, then prior spatial alignments do not convey direct information about the alignment of I_1^k and I_2^k. However, the modality relationship is usually assumed to be invariant along the sequence and hence information about the modality relationship gained from the prior alignments is potentially useful in the registration of the image pair (I_1^k, I_2^k).

An image registration method that does not use prior information gained from previous alignments will be called a *blind* method. Our goal is to study how prior information obtained from previous registration of multi-modal images or volumes can be used to help in the registration problem.

The problem of using prior information to improve multi-modal registration performance was first suggested by Leventon et al. [9]. They propose estimating the underlying joint prior intensity distribution of registered image pairs using training data and then employing a maximum likelihood approach to define the registration measure for new image pairs. Subsequently, Chung et al, [2], proposed an alternative approach in which the quality of registration is determined by the Kullback-Leibler divergence between the estimated joint intensity distribution of pre-aligned data and the joint intensity distribution of the new images. Registration is then accomplished by minimizing this K-L divergence. Both [9] and [2] indicate experimentally that using prior information produces a registration function with a wider basin of attraction, making the algorithm more robust to bad initializations, and a registration algorithm that is faster compared to competing methods.

Our main contribution is to incorporate the use of prior information, as explored in [9] and [2], into a graph theoretic image registration framework. In this approach, we employ entropic spanning graphs [11] to define the registration function and propose a new multi-modal image registration algorithm that incorporates knowledge from previously aligned image pairs. These may be image pairs earlier in a sequence of registration problems or may come from a set of training examples. In either case, as the quality of the prior information improves, the algorithm can use this in a weighted fashion to improve the accuracy of new registrations (of the same class) with minor additional computation.

The remainder of the paper is organized as follows: in Section 2, we provide some background information on information-theoretic registration techniques. Section 3 details the proposed similarity measure, Section 4 reviews the graph theoretic entropy estimation technique employed in this study. In Section 5, we explain the proposed registration algorithm. Section 6 contains experimental results and a discussion.

2. INFORMATION-THEORETIC REGISTRATION METHODS

In simplest form, information theoretic registration methods attempt to measure the spatial alignment of two images or volumes by computing the joint entropy (or a related quantity) of the current corresponding pixels modelled as i.i.d. random variables. Under suitable assumptions, this form of measure can be theoretically justified using Fano's inequality, [1], or its -Renyi generalization [7], [14]. Two methods have been proposed in the literature for computing such a measure. First, for the current alignment one can estimate the mutual information of the pairs of pixels using a so-called "plug-in" estimator [4, 18, 13]. This is based on estimating the density of the data and plugging this estimate in the mutual information formula. Alternatively one can estimate the -Renyi entropy of the current alignment using an entropic spanning graph [11], [12], [16]. As argued in [16], the joint -Renyi entropy of the pixel intensity values can be used as a measure of dissimilarity between two images.

Our study is based on using an entropic spanning graph, in particular a minimum spanning tree (MST), to compute a registration measure. As we shall see this method provides a natural mechanism for incorporating prior information.

2.1 -Renyi Entropy and Jensen Divergence

The -Renyi entropy of a random variable X with density p_X is:

$$H\ (X) = H\ (p_X) = \frac{1}{\ -1}\log\int p_X(x)\mathrm{dx}.$$

where $\ \in (0,1)$ is a parameter. It is known that as $\ \to 1$, $H\ (X) \to H(X)$ where $H(\cdot)$ is the Shannon entropy [5].

The -Jensen distance is a distance measure between probability distributions defined using the Renyi entropy. For a fixed $\ \in (0,1)$, $w \in [0,1]$ the -Jensen distance from p_X to p_Y is:

$$\begin{aligned} J_{\ ,w}(p_X,p_Y) &= H\ (wp_X + (1-w)p_Y) \\ &\quad -[wH\ (p_X) + (1-w)H\ (p_Y)]. \end{aligned}$$

Since H is concave, $J_{\ ,w}(p_X,p_Y) > 0$ when $p_X \neq p_Y$ and $J_{\ ,w}(p_X,p_Y) = 0$ when $p_X = p_Y$ (a.e.).

The -Jensen distance has previously been proposed as a blind registration measure, [10, 8]. In [8] it is used to measure the distance between the observed conditional distributions. In [10] it is used simply to measure the distance between the pixel value distributions of the two images in the overlap region.

Inspired by the approach of [2], we suggest that this distance can also be used to measure the quality of alignment based on the discrepancy between the observed joint intensity distribution and a prior distribution.

Let $I_1^*(x,y)$ and $I_2^*(x,y)$ denote two aligned training images from different modalities. For a given test transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, let $I_1(x,y)$ and $I_2^T(x,y) = I_2(T(x,y))$ be two observed images from the same respective modalities. Assume that each pixel intensity value in the image pairs $I^* = (I_1^*, I_2^*)$ and $I^T = (I_1, I_2^T)$ is an independent sample from the distributions p^* and p_T, respectively. Then the distance between these distributions is a useful way of determining the quality of the current alignment. In particular,

$$J_{\ ,w}(p^*, p_T) \tag{1}$$

can be employed as a registration measure that incorporates prior training data.

3. A HYBRID MEASURE

The Jensen divergence measure defined by (1) is based on pre-aligned training images. Hence the performance of a registration algorithm based on this measure will depend on the quality and amount of this prior information. On the other hand, the Renyi entropy measure $H\ (p_T)$ evaluates the quality of alignment based only on the observed distribution[16]. We propose to combine these two measures in a hybrid registration measure. Let $\mathscr{I}^*$ and $\mathscr{I}^T$ denote the set of samples from I^* and I^T. Then define the hybrid measure by:

$$Q\ (I_1, I_2^T) = J_{\ ,w}(p^*, p_T) + \ * H\ (p_T), \tag{2}$$

where

$$= |\mathscr{I}^T| / (|\mathscr{I}^*| + |\mathscr{I}^T|) \tag{3}$$

is the weight that determines the relative influence of the Renyi entropy term and $w = 1 -\ $. Note that $|.|$ denotes set cardinality. By using this weight, the influences of individual terms are adjusted automatically proportional to the amount of available prior information, i.e., the size of the training set. It is easy to show that $Q(I_1, I_2^T) = H\ (p_T)$, when there are no training samples, i.e., $\mathscr{I}^* = 0$. However, as the amount of training data increases, the weight of the Renyi entropy term will be discounted.

Since $H\ (p^*)$ does not depend on the current alignment, one can easily show that:

$$R\ (I_1, I_2^T) = H\ ((1-\ ) * p^* + \ * p^T) \tag{4}$$

is equivalent to (2) as a registration measure. Thus, the registration problem boils down to determining T^* such that:

$$T^* = \arg\min_{T \in \mathscr{T}} R\ (I_1, I_2^T),$$

where $\mathscr{T}$ is the set of allowed transformations.

4. THE MINIMUM SPANNING TREE ESTIMATOR

In [11] Hero et al. present the following result to estimate the -Renyi entropy of an underlying p.d.f:
Let $Z_n = \{z_1, \ldots, z_n\}$ be n samples (in $\mathbb{R}^d$) drawn from a Lebesgue density p_Z and let

$$W(Z_n) = \ _{e \in G(Z_n)} \|e\|_2^{d(1-\ )}, \tag{5}$$

where $G(Z_n)$ is the list of edges in the EMST of Z_n and $\|.\|_2$ denotes the Euclidean length of an edge. Then:

$$\lim_{n \to} \log\left(\frac{W(Z_n)}{n}\right) = H\ (p_Z) + c \text{ almost surely,}$$

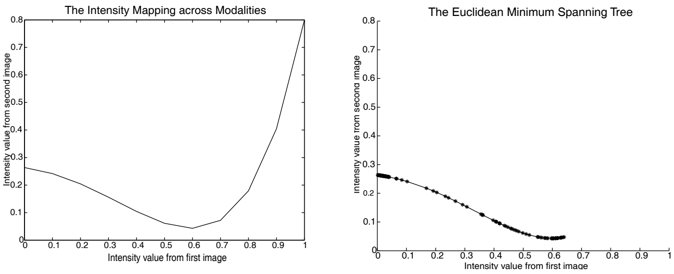
where c is a constant independent of p_Z. This result has been successfully employed to define EMST-based registration measures [12, 15]. One example is to use the total EMST length of pixel intensity samples from $I = (I_1, I_2^T)$ as the objective function. Under the assumption that the samples are i.i.d. the total EMST length corresponds to employing the joint -Renyi entropy as a dissimilarity measure.

As an example, Figure 1-b illustrates a EMST computed over a set of samples from a correctly aligned image pair, where the second image was artificially generated from the first using the intensity mapping function shown in Figure 1-a. We observe that the EMST structure closely follows the cross-modality mapping for the correct alignment.

5. PROPOSED METHOD

We propose to employ an estimate of $R\ (I_1, I_2^T)$ defined in (4) as the registration measure. Note any consistent entropy estimator constructed on $\mathscr{I} = \mathscr{I}^* \cup \mathscr{I}^T$ will converge to (4) as $|\mathscr{I}| \to\ $ and Equation 3 is satisfied. Let $W(\mathscr{I})$, given in (5), denote the total edge length of the EMST computed over the union set of training and observed samples. Then $W(\mathscr{I})$ can be used as a registration function.

A technique to efficiently obtain a descent direction for the fast optimization of a EMST based registration function is given in [16]. A similar argument can be applied to the



(a) (b)

Figure 1: A toy example: a) cross-modality intensity mapping and b) the EMST of samples from a pair of aligned images

measure proposed in this paper: Let the spatial transformation $T(x,y) \in \mathbb{R}^2$ be parameterized by m parameters, i.e., $T_{\mathbf{t}}$ where $\mathbf{t} = (t_1, \ldots, t_m)$. For example, a rigid-body transformation $T_{\mathbf{t}}^R$ has three parameters: two shifts (t_x, t_y) and a rotation and can be expressed as:

$$T_{\mathbf{t}}^R(x,y) = \begin{pmatrix} \cos & \sin \\ -\sin & \cos \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}. \quad (6)$$

Let I_1^* and I_2^* be pre-aligned training images from two different modalities and $\mathscr{I}^* = \{i^* : i^* = (I_1^*(x,y), I_2^*(x,y))\}$. Let I_1 and I_2 be two unregistered images in the same respective modalities. For a given set of transformation parameters, $\mathbf{t}_0$, let $\mathscr{I}^{T_{\mathbf{t}_0}} = \{i : i = (I_1(x,y), I_2(T_{\mathbf{t}_0}(x,y)))\}$. $G(\mathscr{I})$ denotes the set of edges that belong to an EMST of $\mathscr{I} = \mathscr{I}^* \cup \mathscr{I}^{T_{\mathbf{t}_0}}$. For all $e \in G(\mathscr{I})$, let $\|e\|_2$ denote the m-dimensional gradient of the edge length with respect to the transformation parameters. Then, using the main result in [16]:

$$u = - \sum_{e \in G(\mathscr{I})} \|e\|_2 \quad (7)$$

is a descent direction for the registration measure $W(\mathscr{I})$. Hence, u can be used to iteratively update the transformation parameters when searching for the optimum alignment that minimizes $W(\mathscr{I})$.

Consider the two simulated brain images [3] shown in Figure 2. Figure 3 shows two EMST's: EMST-1 of a set of training samples (obtained from a correctly aligned pair of images) and EMST-2 of a set of observed samples (obtained from an image pair misaligned by a 5-pixel translation). Figure 4 shows an EMST of the union set of samples (EMST-3) (training and observed). As discussed in Section 4, EMST-1 is expected to closely follow the cross-modality mapping. EMST-2 and 3 are merely estimates of this mapping and the goal of registration can be viewed as seeking the geometric transformation that best aligns the observed samples with EMST-1. An optimization scheme based on the descent direction u, given in (7), is highly sensitive to the computed EMST structure, i.e., the edges included in the EMST. A bad initialization of the algorithm may lead to an incorrect answer since the transformation updates may be driven by edges that do not "capture" the cross-modality mapping. Training samples, however, are stationary (i.e. don't depend on the tested transformation) and provide an "anchoring" effect (as seen in EMST-3) that helps the EMST structure to follow the cross-modality mapping. This suggests that training samples should make the registration algorithm more robust against bad initialization due this "anchoring" effect.

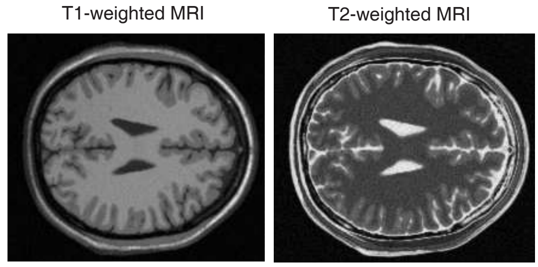


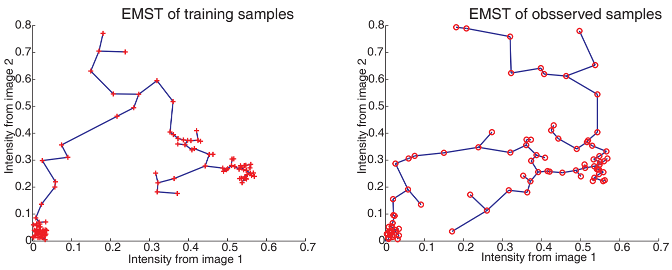
Figure 2: Simulated Magnetic Resonance Images from http://www.bic.mni.mcgill.ca/brainweb/

Adding training samples naturally introduces a computational overhead and slows down the algorithm. This overhead can be minimized using an EMST of the training samples, which can be computed off-line. Based on the following result from [6], all edges that connect two training samples but do not belong to this EMST can then be discarded when computing the registration function: Let $E(\mathscr{I})$ denote the complete set of edges of $\mathscr{I}$ and $G(\mathscr{I})$ denote the corresponding EMST. One can show that:

$$G(\mathscr{I}) \subset ((E(\mathscr{I}) - E(\mathscr{I}^*)) \cup G(\mathscr{I}^*)).$$

6. EXPERIMENTAL RESULTS

The proposed registration algorithm was tested using the simulated MR images shown in Figure 2. For each experiment, the second image was rigidly transformed with a known set of parameters. Thus, the correct alignment was known a priori. Both images were then corrupted with i.i.d zero mean Gaussian noise with a variance equal to 1% of the signal magnitude. The original images were used as the training data-set. Table 1 provides the averaged registration results for three different cases. Registration results from a MI-based implementation [18] are intended to serve as a benchmark. All three algorithms use the same optimization scheme with compatible parameter values. Hence the registration results should provide a fair preliminary comparison of the three registration functions. The first two cases correspond to a relatively large misalignment. The (trained) EMST based algorithm that uses prior knowledge outperforms the other two algorithms in these cases. In general



(a)EMST-1 (b)EMST-2

Figure 3: EMST's of a set of a)training b)observed samples

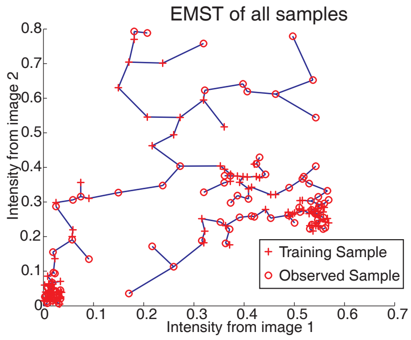


Figure 4: EMST of the union set of samples (EMST-3)

MI has a better accuracy than the (blind) EMST-based algorithm that uses no prior information. The high RMS errors for the two blind algorithms (MI and EMST) are typically due to getting trapped in local minima. The trained EMST algorithm, on the other hand, did not suffer from this problem. The third case corresponds to a small misregistration and all three algorithms achieve sub-pixel accuracy. The blind EMST algorithm yields slightly better registration results than the trained EMST algorithm. This is due to imperfect prior knowledge.

correct	tx = -23	ty = 30	theta = 5
MI	-23.05 (0.92)	30.69 (0.95)	6.81 (2.01)
EMST	-15.75 (13.21)	26.68 (8.71)	3.63 (10.07)
EMST*	-22.88 (0.61)	30.04 (0.24)	4.64 (1.06)
correct	tx = 3	ty = 37	theta = -10
MI	-8.55 (11.56)	33.15 (3.85)	-3.66 (13.71)
EMST	-5.23 (8.31)	11.99 (25.07)	7.86 (17.96)
EMST*	4.97 (1.97)	38.38 (1.38)	-12.16 (2.17)
correct	tx = 5	ty = -2	theta = 0
MI	4.79 (0.21)	-2.06 (0.06)	-0.05 (0.05)
EMST	4.88 (0.12)	-2.19 (0.23)	0.17 (0.18)
EMST*	5.35 (0.40)	-2.07 (0.35)	0.47 (0.67)

Table 1: Avg. registration results (RMS error). MI: Mutual Information based algorithm, EMST = EMST-based algorithm with no prior knowledge, EMST* : EMST-based algorithm with prior knowledge, = 0.5 (averaged over 30 trials)

7. DISCUSSION

The notion of incorporating prior knowledge to image registration has been investigated in the literature. These studies have yielded promising algorithms that achieve a performance better than the standard algorithms that don't employ prior knowledge. In this paper, we proposed and investigated a method to incorporate prior knowledge into a graph-theoretic registration framework. The proposed method achieves this in a natural way with minimal computational overhead. Initial experimental results are encouraging and suggest the algorithm requires further investigation.

REFERENCES

[1] T. Butz, O. Cuisenaire, and J.P. Thiran, "Multi-modal medical image registration: From information theory to optimization objective," *14th Inter. Con. on DSP*, 2002.

[2] A.C.S. Chung, W.M. Wells, A. Norbash, W.E.L Grimson "Multi-modal image registration by minimising Kullback-Leibler distance," *Proc. of MICCAI'02*, Berlin Heidelberg 2002.

[3] C.A. Cocosco, V. Kollokian, R.K.S. Kwan, A.C. Evans, "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database" *NeuroImage*, vol. **5**, no. **4**, 1997.

[4] F. Maes, A. Colligon, D. Vandermeulen, G. Marchal, P. Suetens: "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. **16**, no. **2**, 187–198, 1997.

[5] T.M. Cover and J.A. Thomas, *Elements of Informtion Theory,* Wiley Series in Telecommunications, 1991.

[6] D. Eppstein, "Spanning trees and spanners," *Tech. Report 96-16*, Dept. of Info. and Com. Sci., Uni. of California, Irvine, May 1996.

[7] D. Erdogmus, J.C. Principe, "Information transfer through classifiers and its relation to probability of error" *Proc. IJCNN '01.*, Wahington DC, 2001.

[8] Y.He, A.B. Hamza, and A.H. Krim "Information divergence measure for ISAR image registration" *Proc. of SPIE*, vol. **4379**, pp. 199–208, October 2001.

[9] M.E. Leventon, W.E.L. Grimson "Multi-modal volume registration using joint intensity distribution," *Proc. of MICCAI '98*, Berlin Heidelberg 1998.

[10] A. O. Hero, B. Ma, O. Michel, and J. D. Gorman, "Alpha-divergence for classification, indexing and retrieval, *Technical Report 328*, Comm. and Sig. Proc. Lab. (CSPL), Dept. EECS, University of Michigan, Ann Arbor, July, 2001.

[11] A.O. Hero, B. Ma, O. Michel, J.D. Gorman, "Applications of entropic spanning graphs," *IEEE Signal Proc. Mag.*, vol **19**, no **5**, pp. 85–95, 2002.

[12] H.F. Neemuchwala and A.O. Hero "Entropic graphs for registration" *Multi-sensor image fusion and its applications*, Marcel-Dekker, Inc 2004.

[13] J.P.W. Pluim, J.B.A. Maintz and M.A. Viergever, "Mutual Information based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, 2003.

[14] M.R. Sabuncu, P.J. Ramadge, "Spatial Information in Entropy based Image Registration." *International Workshop on Biomedical Image Registration '03*, Philadelphia, July 2003.

[15] M.R. Sabuncu, P.J. Ramadge "Gradient based nonuniform subsampling for information-theoretic alignment methods," *Proc. of EMBC'04*,San Francisco, Sept 2004.

[16] M.R. Sabuncu, P.J. Ramadge "Gradient based optimization of an EMST image registration function," *Proc. of ICASSP '05*, Philadelphia, March 2005.

[17] J. M. Steele, "Probability theory and combinatorial optimization," *CBMF-NSF regional conf. in appl. math*, vol. **69**, SIAM, 1997.

[18] P.A. Viola, W.M. Wells III, H. Atsumi, S. Nakajima, R. Kikinis, "Multi-modal Volume Registration by Maximization of Mutual Information," *Medical Image Analysis*, vol. **1**, no. **1**, pp. 5–51, 1996.