FAST ALIGNMENT OF DIGITAL IMAGES USING A LOWER BOUND ON AN ENTROPY METRIC

Mert R. Sabuncu, Peter J. Ramadge

Princeton University
Department of Electrical Engineering
Princeton, NJ 08544

ABSTRACT

We propose a registration algorithm based on successively refined quantization and an alignment metric derived from a minimal spanning tree entropy estimate. The metric favors edge alignment, is fast to compute, and compares well in experiments with competing approaches.

1. INTRODUCTION

Registering images is a fundamental problem in image processing. The goal is to align physical or physiological correspondences in two images by minimizing a misalignment metric. Information-theoretic metrics, [11, 10], are based on treating pixel intensity values (or other features) as samples of random variables and using entropy (or related quantities) as the measure of alignment. This paper is based on this approach and develops a fast method for aligning 1-D and 2-D signals based on successively refined quantization and an alignment metric derived from the estimation of entropy via a minimal spanning tree.

The intuitive idea behind information-theoretic metrics is that the dependency between the random variables from two images is maximized at perfect alignment. This dependency can be measured using, for example, mutual information (MI) [11, 1], joint Renyi entropy [5] and Renyi divergence [2], etc. Such metrics have the advantage of permitting direct registration of multi-modal images obtained from different imaging techniques, e.g. CT, MR, etc.

There have been various studies on entropy based registration, e.g. see the references in [10]. At this point, there is no agreement on the best metric to use in a given situation and there are competing methods for computing entropy metrics from the image data. For example, so-called "plug-in" methods [3] are widely used since they are fast to compute and can provide gradient information. However, they are unsuitable when the feature space has a high dimension. An alternative method computes a Euclidean Minimum Spanning Tree (EMST) to estimate entropy [5]. This is equally applicable in higher dimensions but it is relatively slow to compute and until recently it was not clear how to obtain gradient information [9].

Traditional entropy metrics typically use only pixel intensity information. This ignores potentially important information such as spatial location and spatial features. For example, if image A is related to image B by an unknown 1:1 point intensity mapping, then edge locations are an invariant of the modality change and hence should be a useful feature for multi-modal registration. Various prior work has recognized the advantages of incorporating such additional features, e.g. spatial [6, 7, 4, 8] and gradient [6] information. However, this also places an additional burden on the computational resources. When using only pixel intensities, for example, entropy metrics computed via EMSTs are already memory and time intensive. As a result, images are commonly uniformly sub-sampled to reduce image dimensions prior to registration. Alignment metrics that are faster to compute are thus of particular interest.

We propose a fast method for the registration of images motivated by an EMST computation of an entropy metric. There are two novel aspects to the proposed method. First, instead of subsampling the images, we use successively refined quantization to aggregate image information. The quantization produces images that are "step-like". By closely examining the computation of EMSTs for aligning step signals, we show how to lower bound and closely approximate the EMST weight without actually computing the EMST. This provides a second computational advantage. Moreover, the method naturally exploits the alignment of edges in the quantized signals, something that intuition suggests is valuable.

In Section 2 we consider the translation alignment of 1-D step signals using sampling and the computation of EMST's. The main results are a useful lower bound on the EMST weight and showing that the best alignment occurs when edges are aligned. In Sections 3 and 4 we then show how these ideas can be applied to the registration of practi-

Thanks to C.S.B.M.B at Princeton University for providing the fMRI images and **S**iemens **C**orporate **R**esearch, Princeton for a summer internship.

cal signals and images, e.g. medical MR images, by using successively refined quantization. The proposed method is then compared experimentally to the common technique of computing a EMST for uniformly sub-sampled images in Section 5. Under the constraint of commensurate computation time and memory allocation, we show that the proposed algorithm yields a computed registration metric that is smoother, has fewer local minima, a sharper peak and a significant minimum at the correct alignment.

2. 1-D STEP FUNCTIONS & EMST WEIGHT

As it will be illustrated, EMST's defined on samples taken from signals can be employed for the alignment problem. Before going into the details we present some fundamental definitions and a useful property of EMST's. An EMST of a set of points $P \subseteq \mathbb{R}^k$ is a minimum spanning tree of the complete graph with vertex set P and edge weights given by Euclidean distance. The weight of any EMST defined over P is denoted by $W(P)$.

Lemma 1 *Let $V = \{v_j\}$ be a set of n vertices and $z\colon V \to \mathbb{R}^k$ assign each vertex a point in $\mathbb{R}^k$. Let $P = \{z(v_j)\}$ be the image of z and let $W_V(z)$ denote the weight of a MST of the complete graph with vertex set V and edge weights $w_{i,j} = \|z(v_i) - z(v_j)\|$. Then $W_V(z) = W(P)$.*

By Lemma 1 the MST weight is invariant if we sample the vertices keeping only one element for each point value in $\mathbb{R}^k$. This property yields a useful simplification in the computation of EMST's for pairs of step functions.

A function $x(t)\colon [0,T] \to \mathbb{R}$ is a step function if it takes only a finite number of distinct values and has a finite number of discontinuities. For a step function $x(t)$ on $[0,T]$, let $J_x = \{d_i, i = 1 \dots N_x\}$, with $0 < d_1 < \dots < d_{N_x} < T$, be the set of discontinuity points. Set $d_0 = 0$, $d_{N_x+1} = T$, and $\Delta_x = \min_{i=0,\dots,N_x} |d_{i+1} - d_i|$.

These definitions have obvious extensions to vector valued functions: $z\colon [0,t] \to \mathbb{R}^k$ is a step function if each component is a step function. If $z(t) = (x(t), y(t))$ is a step function on $[0,T]$, then $J_z = J_x \cup J_y$. The points in J_z can be ordered and Δ_z defined as above.

To compute the entropy alignment metric of two step functions x and y defined on $[0,T]$, we compute the EMST weight of M samples of the vector signal $z = (x,y)$. More generally, a MST could be employed with (Euclidean distance)$^\gamma$ used as edge weights. This is known to yield an estimate of the joint α-Renyi Entropy, where $\alpha = (2-\gamma)/2$ [4]. We say that $\{t_i, i = 0, \dots N_z\}$ is a *critical set of sample times* for z if $t_i \in [d_i, d_{i+1})$, $i = 0, \dots, N_z$. The corresponding set of samples $\{z(t_i), i = 0, ..., N_z\}$ is said to be a *critical set of samples*. By Lemma 1, the EMST weight of any set of critical samples is the same. Denote this weight by $W_*(z)$. The next lemma indicates that the same is true of any set of sufficiently dense (supercritical) set of samples.

Lemma 2 *For step function $z(t) = (x(t), y(t))$, let $S = \{t_k\}$ be a set of sampling times such that for each $i = 0, ..., N_z$ there exists $t_k \in [d_i, d_{i+1})$. Then $W_S(z) = W_*(z)$.*

Obviously the conditions of the above lemma hold for uniform sampling provided the sample interval T_s satisfies $0 < T_s < \Delta_z$. However, the lemma indicates that it is sufficient to sample right after 0 and thereafter just after every discontinuity.

A step function $x(t)$ on $[0,T]$ is L-step if for $d \in J_x$, $|x(d^+) - x(d^-)| = L$. Similarly, a vector valued step function is L-step if each component is L-step. The following lemma provides tight bounds on the EMST weight of a critically sampled L-step function $z(t) = (x(t), y(t))$.

Lemma 3 *Let $z(t) = (x(t), y(t))$ be an L-step function on $[0,T]$, P be the image of a critical set of samples from z, $E = J_x \cap J_y$ and $|.|$ denote set cardinality. Then:*

$$(|P|-1)L \le W_*(z) \le (|P| - 1 - |E|)L + \sqrt{2}|E|L.$$

By the above lemma, when no edges of the L-step signals x and y are aligned, the EMST alignment metric is trivial to compute; it is determined by the number of distinct values appearing in any critical sampling $|P|$. Interestingly, $|P| - 1$ corresponds to the exponentially weighted MST weight as $\gamma \downarrow 0$. This is sometimes used as the basis of an estimate for the Shannon entropy . Our lower bound of the EMST weight for step functions is simply $|P| - 1$ scaled by L.

To align two step signals, $x(t)$ and $y(t)$, using translation, we seek τ that minimizes $C(\tau) = W_*(z_\tau)$ with $z_\tau(t) = (x_\tau(t), y(t))$, where:

$$x_\tau(t) = \begin{cases} x(t+\tau), & \text{for } t+\tau \in [0,T]; \\ x(0), & \text{for } t+\tau < 0; \\ x(T), & \text{for } t+\tau > T. \end{cases} \tag{1}$$

Clearly, if x and y are L-step, then so is x_τ and z_τ. Next we show that $C(\tau)$ assumes its minimum value at a translation where at least one discontinuity from each signal is aligned.

Lemma 4 *Let $z(t) = (x(t), y(t))$ be an L-step function on $[0,T]$ and τ_o be such that $J_{x_{\tau_o}} \cap J_y \neq \emptyset$. Then for $|\tau - \tau_o| < \Delta_{z_{\tau_0}}$, $C(\tau_o) \le C(\tau)$.*

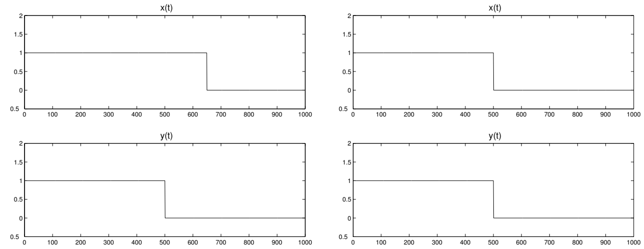
By Lemma 4, at a value τ_o for which some edges in x_{τ_o} and y are aligned, changes of τ around τ_o cannot decrease the value C. Hence there is a local minima of C at τ_o. Lemma 3 shows that around values of τ where edges are not aligned, $C(\tau)$ is a constant. Hence the global minimum value of C will occur at one of the finite number of values of τ where edges are aligned. Thus to align the signals, C need only be computed at such points. A further simplification is possible: instead of computing C, use the lower bound in Lemma 3. Note that $|P|$ will depend on τ. The lower bound is essentially the MST weight in the limiting case $\gamma \downarrow 0$, which corresponds to $\alpha \uparrow 1$. This suggests a connection with Shannon entropy. This idea is explored in the following sections.

3. ALIGNING CONTINUOUS 1-D SIGNALS

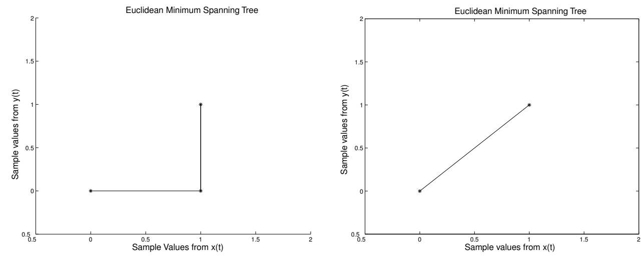
To align two continuous 1-D signals, $x(t)$ and $y(t)$, using translation we first uniformly quantize x and y to q levels. Clearly the quantized signals, x_q and y_q, are L-step functions with L determined by q. This reduces the problem to that of aligning L-step functions. However, a single quantization may result in inaccurate alignment. Hence after the first alignment, increase the value of q, re-quantize and *locally* realign the signals. This can be repeated several times by successively refining the quantization and performing local alignment. Moreover, for each alignment we use the method proposed for the alignment of L-step functions: measure the alignment at translations where edges are aligned using the lower bound given by Lemma 3.

As an example, consider the 1-D signals shown in Fig. 1. The upper graphs show the signals in two possible alignments. The lower graphs show the corresponding EMST of a critical set of samples. When the discontinuities line up, the EMST has a diagonal edge and the number of distinct sample values $|P_{\tau_0}| = 2$ is less than otherwise. When there are no diagonal edges in the EMST the lower bound in Lemma 3 is an equality. By Lemma 3, this also holds when no discontinuities line up. This indicates that it is the proportion of diagonal edges in the EMST that determines how tight the lower bound is.

Figure 2 shows the EMST's for a more complex example. It is clear that only a small fraction of the edges in each EMST are diagonal and hence the lower bound on the EMST weight will be quite tight.

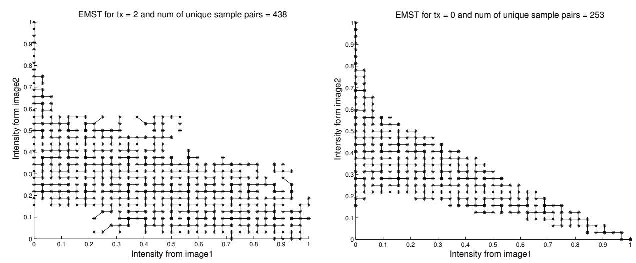


(a) Not Aligned (b) Aligned



(c) Not Aligned (d) Aligned

Fig. 1.



(a) $t_x = 2$ pixels (b) $t_x = 0$ pixels

Fig. 2. EMSTs of two signals for two cases of translation.

Our experiments with a large number of images indicate that for a small number of quantization levels q (e.g. $q \leq 50$) the above lower bound is quite tight.

An algorithm similar to the one discussed in section 3 is proposed to achieve fast image registration. Initially the two images are quantized to a small number of levels and the registration that minimizes $|P|$ (from Lemma 5) is determined. This can then be repeated several times by successively refining the quantization and performing local realignment. For each alignment, it is sufficient to use the definition of a critical set of samples to determine the sampling locations. Typically (especially for q small) the size of this set is much smaller than the total number of possible samples. After fixing the sample locations, the spatial transformation (and its inverse) is applied to these locations to determine the corresponding locations in the other image.

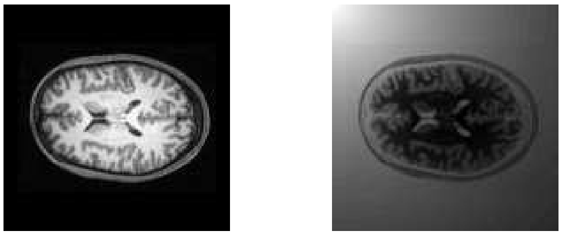
4. REGISTRATION OF IMAGES

We now apply these ideas to image registration. For a quantized digital image U of size $H \times W$, create the vector u by concatenating all the rows of U and let $u(t)$ be the step function obtained by interpolating u using the nearest neighbor rule. For given 2-D images U and V, call the critical set of samples from $z(t) = (u(t), v(t))$ a critical set of samples from (U, V) and denote its EMST weight as $W_*(U, V)$. The following lemma provides a lower bound on $W_*(U, V)$.

Lemma 5 *Let U and V be two digital images with quantization step size L and let P the image of a critical set of samples from (U, V). Then $W_*(U, V) \geq (|P| - 1)L$.*

5. EXPERIMENTAL RESULTS

Space allows only one example. The second image in Figure 3 was generated from the first and is in correct alignment. The results shown below are for translation alignment along the x-axis. However, the proposed method is applicable to other transformations and generates similar results.



(a) Image 1 (b) Image 2

Fig. 3.

Figure 4 shows the proposed metric (normalized to [0,1]) for two quantization levels. Figure 5 shows the proposed metric and the EMST weight. The curves are indistinguishable for 32 quantization levels.

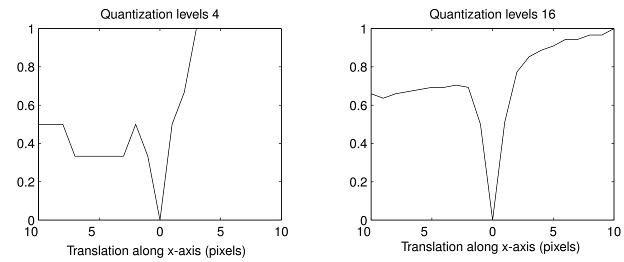
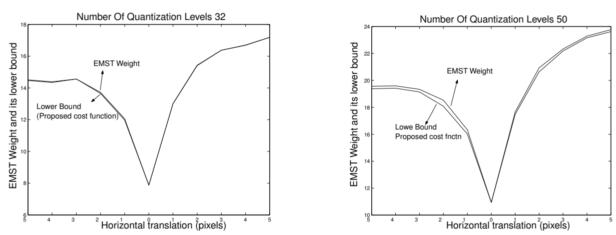


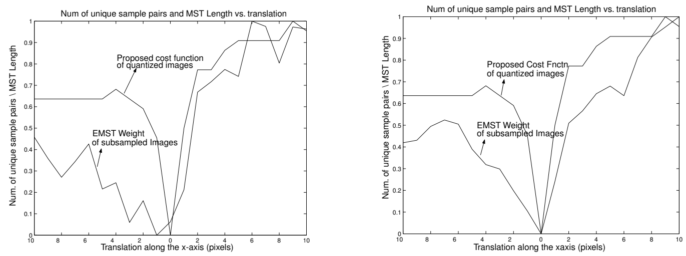
Fig. 4.



(a) 32 Quantization Levels (b) 50 Quantization Levels

Fig. 5.

Figure 6 compares the proposed metric on quantized images and the EMST weight on uniformly subsampled images (both normalized to [0,1]). Note that in both cases the computational time and memory required are much less than for the EMST computation (5-25 times faster in our implementation). The graphs indicate that the proposed metric has fewer local minima, is smoother, and has a sharper minimum at correct transformation, compared to the EMST weight when similar computational resources are allocated.



(a) Subsampling factor=16 (b) Subsampling factor=8

Fig. 6.

6. ACKNOWLEDGEMENTS

The authors thank Prof. Alfred Hero and his group at the University of Michigan, for valuable comments.

7. REFERENCES

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

[2] Y. He, A.B. Hamza, H. Krim, "An Information Divergence Measure For Isar Image Registration," *Proc. IEEE Workshop on Statistical Signal Processing '01*, Singapore, August 2001.

[3] J. Beirlant, E.J. Dudewicz, L. Gyorfi and E.C. van der Meulen. "Nonparametric entropy estimation: an overview," *International Journal Math. Stat. Sci.*, 6, pp. 17-39.

[4] H.F. Neemuchwala, A.O. Hero and P.L. Carson, "Image registration using alpha-entropy measures and entropic graphs," *European Journal on Signal Processing*, 2002.

[5] B. Ma, A.O. Hero, J.D. Gorman and O. Michel, "Image Registration with Minimum Spanning Tree Algorithm," *IEEE International Conf. on Image Processing '00*, vol.1, pp.481-484, Vancouver, BC, Canada, Sept. 2000.

[6] J. Pluim, A. Maintz and M. Viergever "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.*, vol. **19**, no.**8**, pp. 809–814, 2000.

[7] D. Rueckert, M.J. Clarkson, D.L.G. Hill and D.J. Hawkes, "Non-rigid registration using higher-order mutual information," *Medical Imaging: Image Processing*, K.M Hanson, Ed. Bellinghamn, WA, SPIE Press, 2000.

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

[9] M.R. Sabuncu, P.J. Ramadge "Gradient Based Optimization of an EMST Metric for Multi-modal Image Registration." *Technical Report*, Princeton University, June 2004.

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

[11] 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.