Multimodality image registration using an extensible information metric and high dimensional histogramming

Jie Zhang and Anand Rangarajan

Dept. of Computer & Information Science & Engineering,
University of Florida, Gainesville, FL 32611-6120,
{*jiezhang,anand*}*@cise.ufl.edu*

Abstract. We extend an information metric from intermodality (2-image) registration to multimodality (multiple-image) registration so that we can simultaneously register multiple images of different modalities. And we also provide the normalized version of the extensible information metric, which has better performance in high noise situations. Compared to mutual information which can even become negative in the multiple image case, our metric can be easily and naturally extended to multiple images. After using a new technique to efficiently compute high dimensional histograms, the extensible information metric can be efficiently computed even for multiple images. To showcase the new measure, we compare the results of direct multimodality registration using high-dimensional histogramming with repeated intermodality registration. We find that registering 3 images simultaneously with the new metric is more accurate than pair-wise registration on 2D images obtained from synthetic magnetic resonance (MR) proton density (PD), MR T2 and MR T1 3D volumes from Brain Web. We perform the unbiased registration of 5 multimodality images of anatomy, CT, MR PD, T1 and T2 from Visible Human Male Data with the normalized metric and high-dimensional histogramming. Our results demonstrate the efficacy of the metrics and high-dimensional histogramming in affine, multimodality image registration.

1 Introduction

An information *metric* was proposed and used for multimodality image registration in our previous work [1]. Compared to mutual information [2,3,4], the information metric can be easily and naturally extended to multiple random variables and hence can be used to register multiple images simultaneously. (There is no easy and natural corresponding extension for mutual information.) Mutual information of multiple random variables [5] is not necessarily nonnegative, which renders it inadequate as an image similarity measure. Others [6,7,8] have proposed different nonnegative definitions, but they are not natural extensions of the mutual information of two random variables and do not embody the true (in our eyes) spirit of mutual information: shared information between multiple images. Hence, using mutual information to simultaneously register multiple images is not appropriate despite the fact that mutual information is a very good

measure (though not a metric) for registering two images. Our goals in the paper are: first, we wish to perform an unbiased registration of multimodality images and second, we hope to demonstrate that multimodality (multiple images) registration can achieve better accuracy than repeated intermodality (two images) registration.

In order to compute the information metric, we need to estimate the high dimensional probability mass function (PMF) so as to compute the Shannon entropy of multiple random variables. Due to the curse of dimensionality, and especially when derivatives of the PMF are required, it is difficult to accurately estimate a high dimensional probability distribution. The simplest PMF estimation approach is histogramming and has been used in database research such as multiple-attribute-data query [9]. In multimodality image registration, despite the fact that we only need to estimate the entropy (of the form $p \log p$) from the PMF, high dimensional histogramming is not prevalent and due to this, there is almost no previous work on simultaneous, multimodality image registration. In this paper, we use an efficient technique to compute high dimensional histograms so as to efficiently compute the Shannon entropy of multiple images. The technique can compute high dimensional histograms in $O(N)$ time where N is the number of samples (simultaneously drawn from all images). We also show a relationship between the maximum number of bins allowed along each axis of the high dimensional histogram such that the histogram will converge to the true PMF in the high dimensional space as $N \to \infty$.

In the following sections, we describe the information metric and the high dimensional histogramming technique.

2 Multimodality Registration Using The Extensible Metric

Before we move to multimodality registration, we briefly introduce some concepts and the basic registration framework using a metric (technically a pseudometric) in intermodality (2-image) registration. More details about the new metric for intermodality registration can be found in our previous work [1].

2.1 Intermodality Registration Using a Metric

The image similarity metric ρ in the intermodality (2-image) case is the sum of two conditional entropies. For two random variables X and Y,

$$\rho(X,Y) = H(X|Y) + H(Y|X) = 2H(X,Y) - H(X) - H(Y) \tag{1}$$

where $H(\cdot)$ is the entropy of a random variable $[H(X) = -E\{\log(p(X)\}]$, where $p(X)$ is the PMF of X, and $E\{\cdot\}$ denotes the expectation of a random variable. We also proposed two normalized versions of the metric: The first is

$$\tau(X,Y) = \frac{\rho(X,Y)}{H(X,Y)}. \tag{2}$$

$\tau(X,Y)$ is also a pseudometric [10]. And $0 \leq \tau(X,Y) \leq 1$, $\tau(X,Y) = 0$ if $X = Y$; $\tau(X,Y) = 1$ if X and Y are independent. The second normalized version of the

metric $\rho(X, Y)$ is

$$\eta(X, Y) = \frac{\rho(X, Y)}{H(X) + H(Y)}. \tag{3}$$

And $0 \leq \eta(X, Y) \leq 1$, $\eta(X, Y) = 0$ if $X = Y$; $\eta(X, Y) = 1$ if X and Y are independent. But $\eta(X, Y)$ does not satisfy the triangle inequality and hence it is not a metric (or pseudometric).

From the definition in (1), we see that ρ is very similar to mutual information (MI) (4) in the 2-image case except that the metric has one more joint entropy term, which means the metric gives joint entropy more weight than the marginal entropy in comparison to mutual information

$$MI(X, Y) = H(X) + H(Y) - H(X, Y) = \frac{H(X) + H(Y) - \rho(X, Y)}{2}. \tag{4}$$

And we have found that minimizing the normalized metric τ or η is *equivalent* to maximizing the normalized mutual information (NMI) [11] (5) in the 2-image case

$$NMI(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)} = 2 - \tau(X, Y). \tag{5}$$

Consequently, from our perspective, NMI is not *ad hoc* since it is inversely proportional to a pseudometric.

Now we move to our main topic—multimodality image registration.

2.2 Extension to the multimodality case

From the definition of the information metric for two random variables (1), we can *easily* extend the metric to multiple random variables in two different ways. The first extension, for n random variables X_1, X_2, ... X_n,

$$\rho(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i | X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n) \tag{6}$$

and the second is,

$$\mu(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n | X_i). \tag{7}$$

And after dividing by either the joint entropy $H(X_1, X_2, \ldots, X_n)$ or by the sum of the marginals $\sum_{i=1}^{n} H(X_i)$, we get their normalized counterparts.

If we want to simultaneously register three images $I^{(1)}$, $I^{(2)}$ and $I^{(3)}$, we obviously need to find more transformations. We define the biased case as one where $I^{(1)}$ is the reference image and fixed in the registration and we seek two optimal affine transformations—T_2^* for image $I^{(2)}$ and T_3^* for image $I^{(3)}$ by minimizing the metric (8).

$$\{T_2^*, T_3^*\} = \arg \min_{\{T_2, T_3\}} \rho(I^{(1)}, I^{(2)}(T_2), I^{(3)}(T_3)). \tag{8}$$

We define the unbiased case as one where there is no reference image and we seek three optimal affine transformations—T_1^* for image $I^{(1)}$, T_2^* for image $I^{(2)}$ and T_3^* for image $I^{(3)}$ by minimizing the metric (9).

$$\{T_1^*, T_2^*, T_3^*\} = \arg \min_{\{T_1, T_2, T_3\}} \rho(I^{(1)}(T_1), I^{(2)}(T_2), I^{(3)}(T_3)) \tag{9}$$

where $I^{(1)}(T_1)$ is the transformed image of image $I^{(1)}$ using affine transformation T_1, $I^{(2)}(T_2)$ is the transformed image of image $I^{(2)}$ using affine transformation T_2 and $I^{(3)}(T_3)$ is the transformed image of image $I^{(3)}$ using affine transformation T_3. Equivalent minimizations can be carried out for the normalized counterparts of ρ and μ.

3 Computing the entropy of multiple random variables

The multimodality image registration measures ρ, μ and their normalized versions are all entropy-based measures. Consequently, all these measures require the computation of the joint entropy of many random variables—henceforth termed "multi-dimensional entropy."

The approach in [12] used minimum spanning trees (MST) to estimate the α-entropy in image registration. The MST-based approach directly estimates entropy without estimating the high dimensional PMF. But computing an MST for a graph with many edges is very expensive $[O(E \log E)]$ where E is the number of voxels and furthermore, the method cannot compute the normalized versions of the information measure. (Also the method computes the Renyi entropy instead of the Shannon entropy.) Indirect methods compute entropy by first estimating the high dimensional PMF. While histogramming is a popular approach for estimating the PMF, it has not been used for computing the high dimensional entropy in image registration, mainly because *naive implementations are exponential in complexity in the dimensionality of random variables.* Our technique for computing high dimensional histograms (to be explained below) overcomes the aforementioned dimensionality problem. Its computational complexity is $O(N)$ where N is the number of samples drawn from (corresponding) pixel locations over a set of images. The $O(N)$ computational complexity is much smaller than some popular high dimensional PMF estimation methods such as Parzen windows $O(N^2)$, Gaussian mixture models $O(NK)$ where K is the number of clusters, etc. An approximation to the Parzen window entropy can be computed in $O(NM)$, $M < N$ using fast Gauss transforms [13], but you have to first cluster the samples. To our knowledge, this approach has not been explored in medical image registration. Its advantage over the high dimensional histogramming technique is that it is analytically differentiable.

We now describe the high dimensional histogramming approach. Assume we have M images $I^{(m)}$, $m \in \{1, \ldots, M\}$ and the number of histogram bins for the m^{th} image $I^{(m)}$ is $K^{(m)}$, $m \in \{1, \ldots, M\}$. The total number of bins in the multi-dimensional histogram of M images is $\prod_{m=1}^{M} K^{(m)}$, which will be very large if M or $K^{(m)}$ is large. But in the space of the joint histogram of M images, most of the bins of the joint histogram are empty. Empty bins do not contribute anything when we compute the high dimensional Shannon entropy (since $p \log p \to 0$ as

$p \to 0$.) Hence, using $\prod_{m=1}^{M} K^{(m)}$ bins in the space of the joint histogram of M images is impractical and furthermore is unnecessary since we only need know the non-empty bins.

Assume a bounded range $[I_{min}^{(m)}, I_{max}^{(m)}]$ for image $I^{(m)}$. Let $B_i^{(m)}$ be the binned intensity value of image $I^{(m)}$ at location i, $i \in \{1, \ldots, N\}$:

$$B_i^{(m)} = \left[(K^{(m)} - 1) \times \frac{I_i^{(m)} - \min_{1 \le j \le N}\{I_j^{(m)}\}}{\max_{1 \le j \le N}\{I_j^{(m)}\} - \min_{1 \le j \le N}\{I_j^{(m)}\}} + 1 \right], \forall m. \quad (10)$$

From (10), we see that the binned intensity values of image $I^{(m)}$ are integers in $\{1, \ldots, K^{(m)}\}$. Let $L^{(m)}$ be the minimum length of digital bits which can represent $K^{(m)}$, $m \in \{1, \ldots, M\}$. Then we get a new code $C_i = B_i^{(1)} B_i^{(2)} \cdots B_i^{(M)}$ with length $\sum_{m=1}^{M} L^{(m)}$, which is a concatenation of the binned intensity values of *all* images at location i, $i \in \{1, \ldots, N\}$. The number of different elements of the set $\{C_i,\, i \in \{1, \ldots, N\}\}$ is the number of non-empty bins of the joint histogram of M images. Hence we can use $\{C_i,\, i \in \{1, \ldots, N\}\}$ to generate the joint histogram of M images by counting the number of identical C_is in the code set. That this is valid is guaranteed by the following theorem.

Theorem 1: $C_i = C_j$ if and only if $B_i^{(m)} = B_j^{(m)}$ $\forall m \in \{1, \ldots, M\}$, $\forall i, j \in \{1, \ldots, N\}$. [Proof omitted due to lack of space.]

The number of bins $K^{(m)}$ is the only free parameter in our method but it is also very important. Below, following [14], we propose a criterion for limiting the maximum number of bins in the histogram.

Theorem 2: Let $U_1, U_2, \ldots, U_N$ be i.i.d. random variables in $\Re^M$ with PMF f. Let $\mathcal{P}$ be a partition of $\Re^M$ into cubes of size h, and define the histogram PMF estimator by

$$f_N(u) = \frac{1}{Nh^M} \sum_{i=1}^{N} I_{\{U_i \in A(u)\}} \quad (11)$$

where $A(u)$ is the set in $\mathcal{P}$ that contains u and I is the indicator function of a set. Then the estimate is *universally consistent* in L_1 if $h \to 0$ and $Nh^M \to \infty$ as $N \to \infty$, that is, for any f the L_1 error of the estimate $\int |f_N(u) - f(u)|\, du$ converges to zero in probability, or equivalently, for any $\epsilon > 0$,

$$\lim_{N \to \infty} \Pr\left(\int |f_N(u) - f(u)|\, du \ge \epsilon \right) = 0. \quad (12)$$

For our case, the domain of PMF is a bounded subset of $\Re^M$, namely $[0, 1]^M$. If we use the same number $K^{(m)} = K$ bins for each image in the set, then $h = \frac{1}{K}$. Thus $h \to 0$ is equivalent to $K \to \infty$ and $Nh^M \to \infty$ is equivalent to $\frac{N}{K^M} \to \infty$, for which $K < N^{\frac{1}{M}}$ is necessary. Let $K = N^{\frac{1}{M+\alpha}}$, then as $K \to \infty$, $\frac{N}{K^M} = N^{\frac{\alpha}{M+\alpha}} \to \infty$ for any $\alpha > 0$ as $N \to \infty$, which satisfies the condition of the theorem. Hence we use $K = N^{\frac{1}{M+\alpha}}$ for some $\alpha > 0$ as a criterion in our high dimensional histogram. In plain English, Theorem 2 essentially says that if you have more samples, then use more bins for the histogram but the rate of

increase of the number of bins should be slower than the rate of increase of the number of samples. (The simplification of $K^{(m)} = K$ has been used for the sake of exposition. An extension to different $K^{(m)}$ is straightforward.)

From Theorem 1, we know that to compute the high dimensional histogram, we only need to count the number of identical C_i in the set $\{C_i,\, i \in \{1, \ldots, N\}\}$. From Theorem 2, we know that for any C_i, $i \in \{1, \ldots, N\}$, $C_i \in [1, N]$. We can count the number of identical C_i in the set $\{C_i,\, i \in \{1, \ldots, N\}\}$ by traversing N samples once. Thus the time complexity of computing high dimensional histograms is $O(N)$.

4 Experimental Results

4.1 *Multimodality vs. intermodality: Simultaneous registration of 3 images and pair-wise registration on synthetic PD, T2 and T1 MR images*

In the registration experiments of this section, we use the powerful Brainweb simulated MRI volumes for a normal brain [15]. The main advantage of using simulated MR data is that the ground truth is known. The size of each image is 256mm×256mm.

We decompose an affine transformation matrix into a product of shear, scale and rotations. Let $T = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{bmatrix}$ be an affine transformation. $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2^s & 0 \\ 0 & 2^s \end{bmatrix} R(\theta) \begin{bmatrix} 2^t & 0 \\ 0 & 2^{-t} \end{bmatrix} R(\phi)$, where s and t are scale and shear parameters, and $R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$, $R(\phi) = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$ are two rotation matrices. In our experiments, the range of shear and scale parameters is [-1 1], the range of rotation parameters are [-45 45] degrees and the range of translations is [-10 10] mm. In this decomposition of an affine transformation, reflections are not allowed.

In the experiments, we use the following ten measures to register 3 triplets of 2D slices of 3D PD, T2 and T1 MR brain volume images.

1. $\rho(X, Y, Z) = H(X|Y, Z) + H(Y|X, Z) + H(Z|X, Y)$
2. $\mu(X, Y, Z) = H(X, Y|Z) + H(Y, Z|X) + H(Z, X|Y)$
3. $\tau(X, Y, Z) = \frac{\rho(X,Y,Z)}{H(X,Y,Z)}$
4. $\eta(X, Y, Z) = \frac{\rho(X,Y,Z)}{H(X)+H(Y)+H(Z)}$
5. $\kappa(X, Y, Z) = \frac{\mu(X,Y,Z)}{H(X,Y,Z)}$
6. $\sigma(X, Y, Z) = \frac{\mu(X,Y,Z)}{H(X)+H(Y)+H(Z)}$
7. modified mutual information: $mMI(X, Y, Z) = H(X) + H(Y) + H(Y) - H(X, Y, Z)$
8. sum of pair-wise mutual information: $pMI(X, Y, Z) = MI(X, Y)+MI(Y, Z)+MI(Z, X)$
9. modified normalized information: $mNMI(X, Y, Z) = \frac{H(X)+H(Y)+H(Y)}{H(X,Y,Z)}$

10. sum of pair-wise normalized mutual information: $pNMI(X,Y,Z) = NMI(X,Y) + NMI(Y,Z) + NMI(Z,X)$

The slices are chosen in the axial direction. We then transform the MR T2 image with an affine transformation $\hat{T}_1 = \begin{bmatrix} 1.2496 & -0.39666 & 0 \\ 0.39666 & 0.93006 & 0 \\ 4 & 4 & 1 \end{bmatrix}$, with $s_1 = 0.2$, $t_1 = 0.2$, $\theta_1 = 10$, $\phi_1 = 10$, $e_1 = 4$ and $f_1 = 4$. We also transform the MR T1 image with an affine transformation $\hat{T}_2 = \begin{bmatrix} 1.4205 & -0.88097 & 0 \\ 0.88097 & 0.67935 & 0 \\ 8 & 8 & 1 \end{bmatrix}$, with $s_2 = 0.4$, $t_2 = 0.4$, $\theta_2 = 20$, $\phi_2 = 20$, $e_2 = 8$ and $f_2 = 8$. We have done two experiments with these data. Each experiment is repeated 30 times with different Gaussian noise. We add Gaussian noise with zero mean and standard deviation 0.1 in the first experiment and zero mean and standard deviation 0.2 in the second experiment. (The intensity range of all images is normalized to the [0,1] interval.) We use a coarse-to-fine brute force search strategy to find the optimal T_1^* and T_2^*. The finest search resolution of scale and shear is 0.05. The finest search resolution of rotation is 0.5 degrees in the first experiment and 1 degree in the second experiment. The finest search resolution of translation is 1 mm. The registration measures are computed only in the overlap area of the three images with bilinear interpolation used for transforming the image intensities.

To compare each measure and validate the registration results, we compute the mean error of each parameter of two affine transformations recovered by ten measures and the sum of $\left\|\hat{T}_1 - T_1^*\right\| + \left\|\hat{T}_2 - T_2^*\right\|$ of 30 experiments.

Figure 1 depicts the sum of $\left\|\hat{T}_1 - T_1^*\right\| + \left\|\hat{T}_2 - T_2^*\right\|$ of 30 noise trials in the first set of experiments. From the results, we see that multimodality registration is more accurate than repeated pair-wise (intermodality) registration. And the normalized metric η has best performance. Table 1 shows the mean error of each parameter of two affine transformations recovered with the ten measures of 30 noise trials in the second experiment. Figure 2 shows the sum of $\left\|\hat{T}_1 - T_1^*\right\| + \left\|\hat{T}_2 - T_2^*\right\|$ of 30 noise trials in the second experiment. From the results, we see that the two non-normalized metrics ρ and μ failed in recovering scale because of high noise. But the normalized metric η (which is based on ρ) still has best performance. And multimodality registration is more accurate than repeated pair-wise registration.

With these experiments on synthetic PD, T2 and T1 MR 2D images, we see that these ten measures have similar performance in low noise experiments. Generally, they can correctly recover scale and shear parameters but have error in recovering rotation and translation. And we see that multimodality registration is more accurate than repeated pair-wise registration. In the high noise case, the two non-normalized metrics failed to recover scale because they prefer small overlaps of images. The normalized metric η still has best performance. And

these experimental results also show that minimizing the normalized metric κ or σ is equivalent to maximizing the modified normalized mutual information.

In our following experiments, we will only use the normalized measure η.

4.2 *Unbiased multiple image registration of Visible Human Data*

Algorithm 1 Iterated sequential search

1. sequentially search ten translations which minimize η corresponding to 5 images;
2. sequentially search ten scaling and shear parameters which minimize η for 5 images;
3. sequentially search ten rotations which minimize η for 5 images;
4. if η decreases in this iteration then go to 1; else end.

In this section, we register pentads of images of head slice from Visible Human Male Data. For the pentad, the first image is the photograph of anatomical slice, the second is the CT image, the third is an MR PD image, the fourth is an MR T1 image and the fifth is an MR T2 image. The slice number of the pentads in the Visible Human Male Data is 1165. Because ground truth is unknown, we register 5 images simultaneously without bias. That means that each image gets an affine transformation and we minimize the normalized metric η on five affine transformations:

$$\{T_1^*, T_2^*, T_3^*, T_4^*, T_5^*\} = \arg\min_{T} \eta(I^{(1)}(T_1), I^{(2)}(T_2), I^{(3)}(T_3), I^{(4)}(T_4), I^{(5)}(T_5)) \tag{13}$$

where $T = \{T_1, T_2, T_3, T_4, T_5\}$ and T_1, T_2, T_3, T_4 and T_5 are five affine transformations. $I^{(m)}(T_m)$ is the transformed image of image $I^{(m)}$ using affine transformation T_m, $m \in \{1, ..., 5\}$. Since the time complexity of searching for 30 parameters of 5 affine transformations is high, we used iterated sequential search using algorithm 1 for each parameter until the normalized measure η achieves the minimum. The color images of the anatomical slice are converted to grayscale and the intensity of images is normalized to the interval $[0, 1]$ prior to registration. The normalized measure η is computed only in the overlap area of the three images with bilinear interpolation used for transforming the image intensities. The image size is 256 by 256. (The pixel size is 0.32mm square for a photograph of anatomical slice and 1mm for MR and CT images.) The histogram of each image used 8 bins. High dimensional histograms are computed using the technique in section 3.

In Figure 3, we show the images before and after registration. In order for human perception to gauge the results of registration, we add a grid to the images. A careful examination of the images before and after registration reveals that the images are indeed better aligned. For a quantitative evaluation of the registration, we coarsely segment these images by basically segmenting the object from the background in the images. Then we represent these segmented images as binary images as shown in Figure 4. (Object is with intensity value 1 and background is with intensity value 0.) We evaluate the quality of the registration by comparing the number of pixels in the nonoverlap region of pairwise segmented images before and after registration. From these results in Table 3,

we see that the number of pixels in the nonoverlap region of segmented pairwise images after registration is much less prior to registration. Provided that the segmentation errors are not significant and these can also be gauged by human perception, we see the images are better aligned after registration. Also, from the segmented images in Figure 4, we see that these images are better aligned after registration.

From the affine transformations achieved in the registration as shown in Table 2, we see that the affine transformations of all three images include a certain amount of shear. This serves as a very preliminary justification for using an affine mapping.

From these anecdotal evaluation results, we see that minimizing the normalized measure η and computing high dimensional histograms works well for the simultaneous (and unbiased) registration of multimodality images.

5 Conclusions

We have presented an information metric for intermodality image registration, which can be easily extended to the multimodality case as opposed to mutual information which is not so easily extended. The information metric is a linear combination of conditional entropies and has the properties of symmetry, non-negativity and triangle inequality. Normalized versions of this extensible information metric are also proposed and used for multimodality image registration. We derive and use a new efficient technique for computing high dimensional histograms so as to efficiently compute the joint entropy of multiple images. We then demonstrate how the high dimensional histogramming technique can be used to simultaneously register many images without being biased to a reference image. The high dimensional histogramming technique can also be used for feature-based multimodality image registration (where a vector of features is available at each voxel) and for non-rigid multimodality image registration. These represent attractive topics for further research.

Acknowledgements

This work was partially supported by NSF IIS 0307712. We thank Tim Cootes and Derek Hill for helpful conversations.

References

1. Zhang, J., Rangarajan, A.: Affine image registration using a new information metric. In: IEEE Computer Vision and Pattern Recognition (CVPR). Volume 1. (2004) 848–855
2. Viola, P., Wells III, W.M.: Alignment by maximization of mutual information. In: Fifth Intl. Conf. Computer Vision (ICCV), IEEE Press (1995) 16–23
3. Collignon, A., Vandermeulen, D., Suetens, P., Marchal, G.: 3D multi–modality medical image registration using feature space clustering. In Ayache, N., ed.: Computer Vision, Virtual Reality and Robotics in Medicine. Volume 905 of Lecture Notes in Computer Science., Springer–Verlag (1995)
4. Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A.: Mutual-information-based registration of medical images: A survey. IEEE Trans. on Medical Imaging **22** (2003) 986–1004

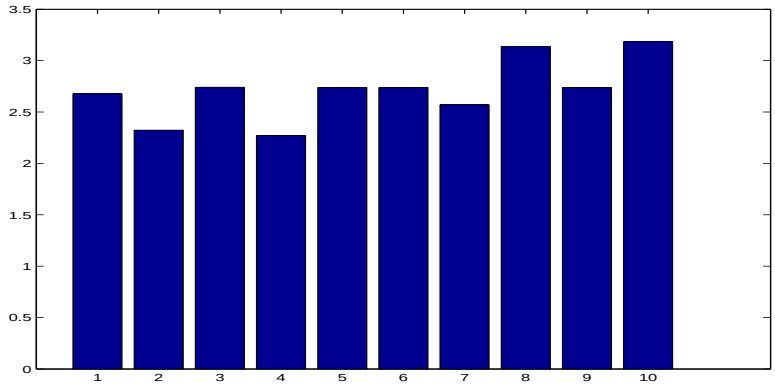


Fig. 1. Plots of sum of $\left\|\hat{T}_1 - T_1^*\right\| + \left\|\hat{T}_2 - T_2^*\right\|$ of 30 trials recovered by ten measures in the first experiment with noise std. 0.1. Numbers 1 to 10 represent ρ, μ, τ, η, κ, σ, mMI, pMI, mNMI and pNMI—the ten registration measures.

Table 1. Mean errors of different affine parameters in the second experiment with Gaussian noise with mean 0 and std. 0.2

	error of transformation on MR T2						error of transformation on MR T1					
measures	s	t	θ	ϕ	e	f	s	t	θ	ϕ	e	f
ρ	0.24	0.01	2.53	1.53	2.53	0.7	0.11	0.02	3.77	2.4	1.73	1.2
μ	0.64	0.06	10.57	3.03	5.93	1.57	0.42	0.08	4.87	3.53	2.97	2.03
τ	0	0	1.1	1.33	1.93	0.43	0.0033	0.0083	3.83	1.67	1.833	0.97
η	0	0	0.97	1.33	1.93	0.33	0.0017	0.0083	2.8	1.23	1.77	0.93
κ	0	0	1.03	1.2	2.07	0.37	0.005	0.0117	3.53	1.43	2.13	0.7
σ	0	0	1.03	1.2	2.07	0.37	0.005	0.0117	3.53	1.43	2.13	0.7
mMI	0	0	1.1	1.2	2.03	0.37	0.005	0.0117	3.93	1.4	2	0.9
pMI	0	0	1.2	1.37	2.17	0.3	0.005	0.013	7.43	1.8	2	0.67
mNMI	0	0	1.03	1.2	2.07	0.37	0.005	0.0117	3.53	1.43	2.13	0.7
pNMI	0	0	1.16	1.37	2.17	0.3	0.005	0.0133	7.4	1.8	2.23	0.77

Table 2. Results of unbiased registration of anatomical slice, CT, MR PD, T1 and T2 images.[Dataset index: VHD #1165.]

	Anatomy	CT	MR PD	MR T1	MR T2
s	-0.1	0.01	0.02	-0.03	0.07
t	-0.01	0.09	-0.01	0.08	-0.02
θ	-7	3	20	0	0
ϕ	-1	-2	0	0	0
e	4	2	-1	0	0
f	-9	1	0	0	0

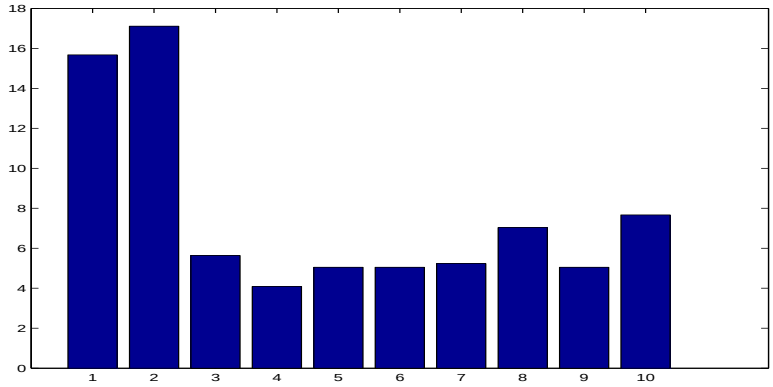
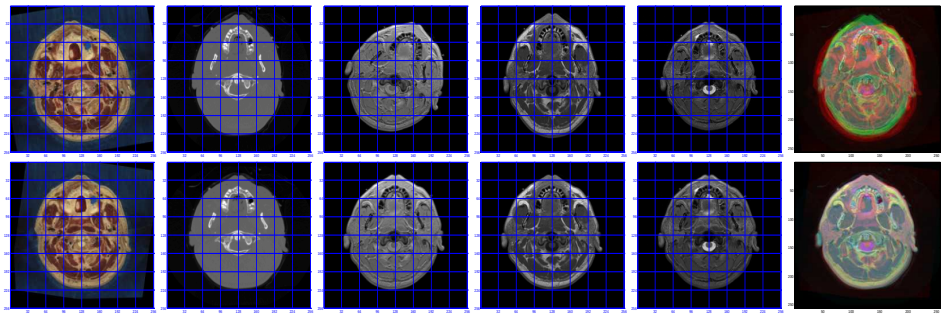


Fig. 2. Plots of sum of $\left\|\hat{T}_1 - T_1^*\right\| + \left\|\hat{T}_2 - T_2^*\right\|$ of 30 trials recovered by ten measures in the second experiment with noise 0.2. Numbers 1 to 10 represent ρ, μ, τ, η, κ, σ, mMI, pMI, mNMI and pNMI—the ten registration measures.

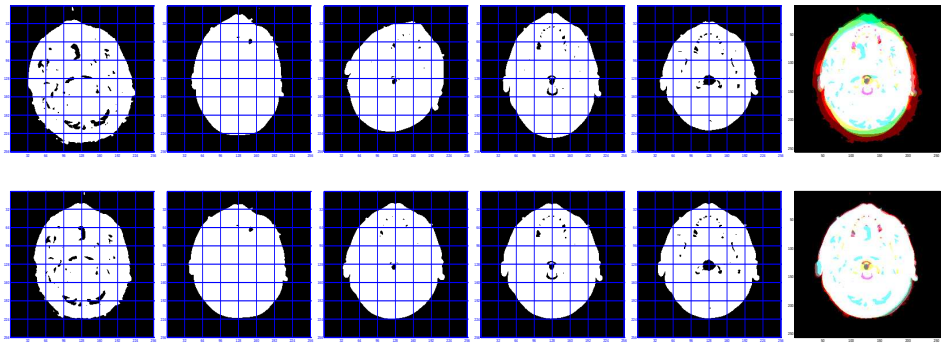


(a) anatomical (b) CT (c) MR PD (d) MR T1 (e) MR T2 (f) overlap

Fig. 3. The first row is the set of images before registration; the second row is the set after registration. [Dataset index: VHD #1165.]

Table 3. Number of pixels in nonoverlap region of segmented images before and after registration. upper triangle is before registration and lower triangle is after registration. [Dataset index: VHD #1165.]

nonoverlap	Anatomic	CT	MR PD	MR T1	MR T2
Anatomic	0	8250	8721	7719	9449
CT	2372	0	4173	1915	3853
MR PD	2984	1530	0	4446	3632
MR T1	2980	1234	1118	0	4292
MR T2	3390	1700	1434	800	0



(a) anatomical (b) CT (c) MR PD (d) MR T1 (e) MR T2 (f) overlap

Fig. 4. Segmented images before (1st row) and after (2nd row) registration. [Dataset index: VHD #1165.]

5. Cover, T., Thomas, J.: Elements of Information Theory. John Wiley and Sons, New York, NY (1991)
6. Studholme, C., Hill, D.L.G., Hawkes, D.J.: Incorporating connected region labelling into automated image registration using mutual information. In Amini, A.A., Bookstein, F.L., Wilson, D.C., eds.: Mathematical methods in biomedical image analysis (MMBIA). IEEE Computer Soc. Press (1996) 23–31
7. Boes, J.L., Meyer, C.R.: Multi-variate mutual information for registration. In Taylor, C., Colchester, A., eds.: Medical image computing and computer-assisted intervention (MICCAI). Volume 1679 of Lecture notes in Computer Science. Springer-Verlag (1999) 606–612
8. Lynch, J.A., Peterfy, C.G., White, D.L., Hawkins, R.A., Genant, H.K.: MRI-SPECT image registration using multiple MR pulse sequences to examine osteoarthritis of the knee. In Hanson, K.M., ed.: Medical Imaging: Image Processing. Volume 3661 of Proc. SPIE. SPIE (1999) 68–77
9. Poosala, V., Ioannidis, Y.: Selectivity estimation without the attribute value independence assumption. In: 23rd VLDB conference. (1997) 486–495
10. Rajski, C.: A metric space of discrete probability distributions. Information and Control **4** (1961) 371–377
11. Studholme, C., Hill, D.L.G., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition **32** (1999) 71–86
12. Neemuchwala, H., Hero, A., Carson, P., Meyer, C.: Local feature matching using entropic graphs. In: IEEE International Symposium on Biomedical Imaging. (2004)
13. Yang, C., Duraiswami, R., Gumerov, N., Davis, L.: Improved fast Gauss transform and efficient kernel density estimation. In: Ninth IEEE International Conference on Computer Vision (ICCV). Volume 1. (2003) 464–471
14. Devroye, L., Gyorfi, L., Lugosi, G.: A Probabilistic Theory of Pattern Recognition. Springer (1997)
15. Collins, D.L., Zijdenbos, A.P., Kollokian, V., Sled, J.G., Kabani, N.J., Holmes, C.J., Evans, A.C.: Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag. **17** (1998) 463–468