Affine Image Registration Using A New Information Metric

Jie Zhang and Anand Rangarajan
Dept. of Computer & Information Science & Engineering
University of Florida, Gainesville, FL 32611
{*jiezhang,anand*}*@cise.ufl.edu*

Abstract

We present a new information metric for multimodality image registration. The metric is technically a pseudometric since it satisfies the properties, i) nonnegativity, ii) symmetry, iii) triangle inequality and is iv) zero if (but not only if) the two image intensities are identical. Information metrics are rarely used in image registration and notably, the widely used mutual information measure is not a metric. Given images A and B, the metric used here is the sum of the conditional entropies $H(A|B)$ and $H(B|A)$. We show that when compared to mutual information which can even become negative in the multiple image case, it is easier to extend our metric to the registration of multiple images. And, after using an upper bound, we show that the sum of the conditional entropies can be efficiently computed even in the multiple image case. We use the metric to simultaneously register multiple 2D slice images obtained from proton density (PD), magnetic resonance (MR) T2 and MR T1 3D volumes and to match human face images obtained under different illuminations. Our results demonstrate the efficacy of the metric in affine, multiple image registration.

1 Introduction

Despite the pervasive use of the mutual information measure in image registration in general and in medical imaging in particular, the measure is not a metric in that it does not satisfy the triangle inequality. The first goal in this paper is to define a new entropy-based image registration measure which is a metric (technically a pseudometric). Entropy-based image similarity measures have become very popular in image analysis especially in situations where there are significant illumination differences between the images being compared. In the present work, we define a new entropy-based image similarity metric. The new metric $\rho(A, B) = H(A|B) + H(B|A)$ where $H(A|B)$ and $H(B|A)$ are the conditional entropies of random variables A and B respectively. The metric is technically a pseudometric since it satisfies three properties of a metric: i) nonnegativity, ii) symmetry, and iii) triangle inequality but not the fourth—$\rho(A, B) = 0$ if $A = B$ but is also equal to zero if $A = f(B)$.

A second goal in this paper is the simultaneous registration of multiple images. In medical imaging, for example, we frequently need to register at the same time, images of different modalities. And there is a frequent need to perform image registration of natural images acquired under different illumination conditions. While the mutual information measure is popular in intermodality (two images) registration, it has not been widely used in multimodality (more than two images) situations. While the mutual information measure between two random variables is non-negative, this is not true when more than two random variables are involved. The mutual information measure between (more than two) random variables *can even become negative*. We show that there is a natural extension of our information metric to the case of three or more images. However, it turn out that the computational complexity of estimating the conditional entropy can become very expensive since it requires estimating a high dimensional probability mass function. We alleviate this problem by minimizing an upper bound of our information metric which is a metric in its own right.

Since Viola and Wells [1] and Collignon *et al.* [2] used mutual information in image registration, there have been hundreds of papers [3] on image registration using entropy-based criteria and mutual information. Two notable facts emerge after surveying the literature. There is almost no prior work on defining an entropy-based image similarity *metric* as we have done here and there are considerable differences of opinion on the extension of mutual information from the intermodality (two images) case to the multimodality (more than two images) case. In [4], mutual information of three random variables X,Y and Z is defined as $MI(X, Y, Z) = H(X) + H(Y) + H(Z) - H(X, Y) - H(X, Z) - H(Y, Z) + H(X, Y, Z)$. Unfortu-

nately, $MI(X,Y,Z)$ is not necessarily nonnegative, which renders it inadequate as an image similarity measure. In [5] and [6], a different definition is proposed: $MI(X,Y,Z) = H(X) + H(Y) + H(Z) - H(X,Y,Z)$. This definition is nonnegative but it is not a natural extension of the mutual information of two random variables. In [7], three images are registered using yet another different definition of mutual information. However, all these definitions 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. Finally, even if we had a nonnegative and natural definition for multimodality mutual information, computing this high dimensional mutual information is still difficult since it requires the estimation of a high dimensional probability mass function, which is computationally very expensive.

2 Description of the information metric

2.1 Metric Definition

After our brief introduction and review of the different inadequacies of the mutual information measure, we now turn to the definition of our new information metric.

The metric ρ is the sum of two conditional entropies and defined in [4]. For two random variables X and Y,

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

where $H(\cdot)$ is the entropy of a random variable and defined as $H(X) = -E(\log(p(X))$, where $p(X)$ is the probability mass function of X, and $E(\cdot)$ denotes the expectation of a random variable. Hence $H(X|Y) = -E(\log(p(X|Y))$ and $H(Y|X) = -E(\log(p(Y|X))$. If X and Y are discrete random variables, $H(X|Y) = -\sum_{y\in\Omega}\sum_{x\in\Omega} p(x,y)\log(p(x|y))$ and $H(Y|X) = -\sum_{y\in\Omega}\sum_{x\in\Omega} p(x,y)\log(p(y|x))$; If X and Y are continuous random variables, $H(X|Y) = -\int_\Omega\int_\Omega p(x,y)\log(p(x|y))dxdy$ and $H(Y|X) = -\int_\Omega\int_\Omega p(x,y)\log(p(y|x))dxdy$, where $p(x,y)$ is the joint probability mass function of X and Y, and $p(x|y)$ and $p(y|x)$ are conditional probability mass functions of X and Y respectively.

$\rho(X,Y)$ is a pseudometric since it satisfies the following four properties [4].

1. $\rho(X,Y) \geq 0$
2. $\rho(X,Y) = \rho(Y,X)$
3. $\rho(X,Y) = 0$ if $X = Y$. However $\rho(X,Y) = 0$ also if $X = f(Y)$. This is why $\rho(X,Y)$ is not a true metric.

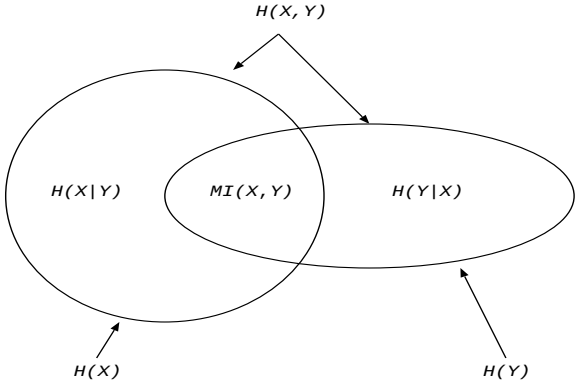


Figure 1: Venn diagram for two random variables

4. $\rho(X,Y) + \rho(Y,Z) \geq \rho(X,Z)$

As explained later, we have also found it useful to work with the following normalized metric. Let

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

Then $\tau(X,Y)$ is also a pseudometric. For proof, please see [8]. Since $\tau(X,Y) = \frac{\rho(X,Y)}{H(X,Y)}$, we henceforth refer to $\tau(X,Y)$ as the normalized metric. 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. Another normalized version of the metric $\rho(X,Y)$ is

$$\begin{aligned} \eta(X,Y) &= \frac{H(X|Y)+H(Y|X)}{H(X)+H(Y)} \\ &= \frac{\rho(X,Y)}{H(X)+H(Y)}. \end{aligned} \tag{3}$$

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).

2.2 Relationship between the information metric and mutual information

The definition of mutual information for two random variables X and Y is

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

where $H(X)$ and $H(Y)$ are marginal entropies of X and Y and $H(X,Y)$ is the joint entropy of X and Y. Also,

$$H(X|Y) = H(X,Y) - H(Y) \tag{5}$$

and

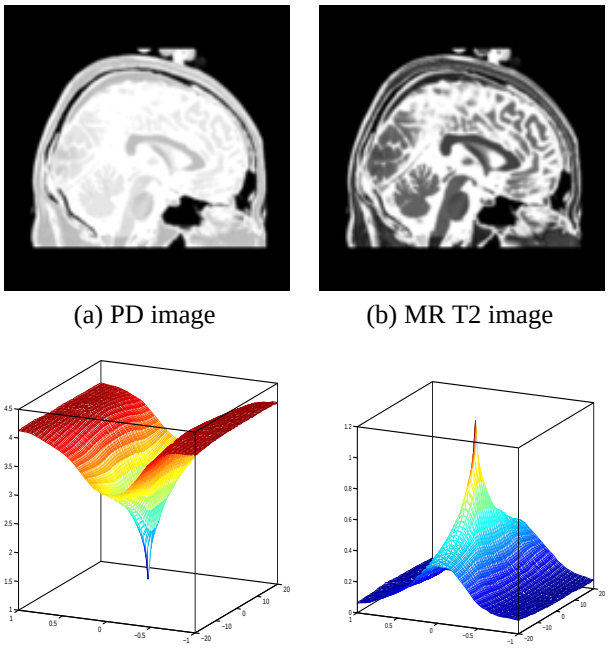
$$H(Y|X) = H(X,Y) - H(X). \tag{6}$$

Hence

$$\begin{aligned}\rho(X,Y) &= H(X|Y)+H(Y|X)\\ &= 2H(X,Y)-H(X)-H(Y)\\ &= H(X)+H(Y)-2MI(X,Y)\\ &= H(X,Y)-MI(X,Y)\end{aligned} \tag{7}$$

From Figure 1, we can visualize the relationship between the metric ρ and MI, where the metric ρ is the non-overlapping region whereas MI is the common region between two sets.

In Figure 2, we plot the values of the metric ρ and MI between a proton density (PD) MR image and a rotated and scaled MR T2 image, where the rotation angle ranges from -20 degrees to 20 degrees and (log) scale ranges from -1 to 1. From Figure 2, the metric ρ is smoothly behaved w.r.t. rotation and scale and achieves its minimum at the point where both rotation and scale are zero; MI is also smoothly behaved w.r.t. rotation and scale and achieves its maximum at the point where both rotation and scale are zero. Hence, at least in this anecdotal example, we can minimize ρ or maximize MI to recover rotation and scale variables.



(a) PD image (b) MR T2 image

(c) the metric ρ (d) Mutual Information

Figure 2: The metric ρ and mutual information between PD and rotated and scaled T2 images

Now we discuss relationships between the normalized metric τ and normalized mutual information (NMI). The definition of normalized mutual information [9] for two random variables X and Y is

$$NMI(X,Y) = \frac{H(X)+H(Y)}{H(X,Y)}. \tag{8}$$

Hence the normalized metric

$$\begin{aligned}\tau(X,Y) &= \frac{H(X|Y)+H(Y|X)}{H(X,Y)}\\ &= \frac{2H(X,Y)-H(X)-H(Y)}{H(X,Y)}\\ &= 2-\frac{H(X)+H(Y)}{H(X,Y)}\\ &= 2-NMI(X,Y)\end{aligned} \tag{9}$$

A different normalized version of the metric

$$\begin{aligned}\eta(X,Y) &= \frac{H(X|Y)+H(Y|X)}{H(X)+H(Y)}\\ &= \frac{2H(X,Y)-H(X)-H(Y)}{H(X)+H(Y)}\\ &= \frac{2H(X,Y)}{H(X)+H(Y)}-1\\ &= \frac{2}{NMI(X,Y)}-1.\end{aligned} \tag{10}$$

Hence, maximizing NMI is equivalent to minimizing the normalized metric τ or η.

2.3 Affine Registration by Minimizing the Metric

Assume that we have two images $I^{(1)}$and $I^{(2)}$. We treat the intensity value of each pixel as an independent random variable. We seek to register image $I^{(2)}$ to image $I^{(1)}$ by determining the best affine transformation T^* which minimizes the metric (1).

$$T^* = \arg\min_T \rho(I^{(1)}, I^{(2)}(T)) \tag{11}$$

where $T = \begin{bmatrix} a & b & 0\\ c & d & 0\\ e & f & 1\end{bmatrix}$ is an affine transformation. In T, the submatrix $\begin{bmatrix} a & b\\ c & d\end{bmatrix}$ can be decomposed into shear, scale and rotation and the vector$\begin{bmatrix} e & f\end{bmatrix}$ contains the x and y translations. The image $I^{(2)}(T)$ is the transformed image of image $I^{(2)}$ using the affine transformation T.

2.4 Extension to the multimodality case

Having defined the metric for two images, we now work out the extension to the multi-modality case. Again, please note that this extension is extremely straightforward as opposed to extending mutual information.

From the definition of the information metric for two random variables (1) and from the Venn diagram shown below in Figure 3, we can easily extend the metric to three random variables. For three random variables X, Y and Z,

$$\begin{aligned}&\rho(X,Y,Z)\\ &= H(X|Y,Z)+H(Y|X,Z)+H(Z|X,Y)\end{aligned} \tag{12}$$

where, as before, $H(\cdot)$ is the entropy of a random variable and defined as $H(X) = -E(\log(p(X))$, with $p(X)$ being the probability mass function of X. Hence $H(X|Y,Z) = -E(\log(p(X|Y,Z))$, $H(Y|X,Z) = -E(\log(p(Y|X,Z))$ and $H(Z|X,Y) = -E(\log(p(Z|X,Y))$.

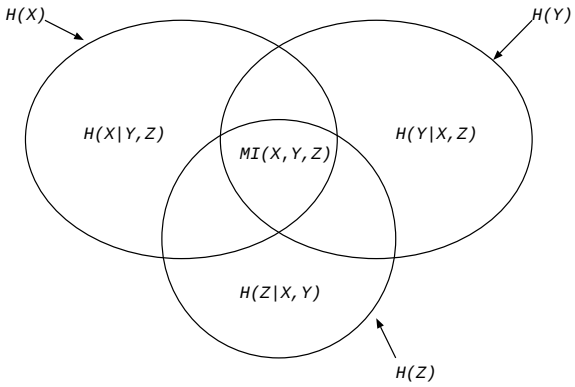


Figure 3: Venn diagram for three random variables

From Figure 3, we can see the relationship between the metric ρ and MI, where the metric ρ is the sum of the entropies of the non-overlapping regions whereas MI is the common region of the three sets.

For n random variables X_1, X_2 ... X_n,

$$\begin{array}{l}\rho(X_1, X_2, \ldots, X_n) \\ = \sum_{i=1}^{n} H(X_i|X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)\end{array} \quad (13)$$

When we simultaneously register three images $I^{(1)}$,$I^{(2)}$ and $I^{(3)}$, we seek two optimal affine transformations T_1^* and T_2^* by minimizing the metric (12).

$$\{T_1^*, T_2^*\} = \arg \min_{\{T_1, T_2\}} \rho(I^{(1)}, I^{(2)}(T_1), I^{(3)}(T_2)) \quad (14)$$

where $T_1$and T_2 are two affine transformations. $I^{(2)}(T_1)$ is the transformed image of image $I^{(2)}$ using affine transformation T_1 and $I^{(3)}(T_2)$ is the transformed image of image $I^{(3)}$ using affine transformation T_2.

We can also extend the normalized metric τ and η to multiple variables:

$$\begin{array}{l}\tau(X_1, X_2, \ldots, X_n) \\ = \frac{\sum_{i=1}^{n} H(X_i|X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n)}{H(X_1,\ldots,X_n)} \\ = \frac{\rho(X_1,X_2,\ldots,X_n)}{H(X_1,\ldots,X_n)}\end{array} \quad (15)$$

$$\begin{array}{l}\eta(X_1, X_2, \ldots, X_n) \\ = \frac{\sum_{i=1}^{n} H(X_i|X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n)}{\sum_{i=1}^{n} H(X_i)} \\ = \frac{\rho(X_1,X_2,\ldots,X_n)}{\sum_{i=1}^{n} H(X_i)}\end{array} \quad (16)$$

For the problem of registering three images, we can use the normalized metric τ or η. The objective function in (14) uses τ or η instead of ρ.

2.5 Computational complexity considerations

From the definition of the metric ρ for three random variables, it can be seen that we have to estimate the joint probability mass function $p(X, Y, Z)$ of three random variables X, Y and Z, which is computationally much more expensive than the estimation of the joint probability mass function of two random variables. To alleviate the computational burden, we present an upper bound of the metric ρ which is also a metric. We then use the upper bound as a legitimate metric for registration instead of ρ. It turns out that the upper bound does not require the estimation of the joint probability mass function of three random variables. This is the main benefit of using the upper bound.

$$\begin{array}{l}\rho(X, Y, Z) \\ = H(X|Y, Z) + H(Y|X, Z) + H(Z|X, Y) \\ \\ \leq \frac{1}{2}(H(X|Y) + H(X|Z)) \\ \\ +\frac{1}{2}(H(Y|X) + H(Y|Z)) \\ \\ +\frac{1}{2}(H(Z|X) + H(Z|Y)) \\ \\ = \frac{1}{2}(\rho(X, Y) + \rho(Y, Z) + \rho(X, Z))\end{array} \quad (17)$$

In the reduction of (17), we mainly use the property of entropy: $H(X|Y, Z) \leq H(X|Y)$.

Let

$$\kappa(X, Y, Z) \stackrel{\text{def}}{=} \frac{1}{2}(\rho(X, Y) + \rho(Y, Z) + \rho(X, Z)) \quad (18)$$

In multimodality image registration, we may minimize (18) instead of minimizing (12).

Finally, we briefly discuss the field of view problem which besets much of entropy-based image registration. Since the joint probability mass function between two images is estimated by considering the overlap region of the two images, the trivial minimization of the information metric and maximization of the mutual information occurs when the two images are spatially separated with zero overlap. The information metric ρ is minimized in that case and the mutual information reaches its maximum value equal to the sum of the two marginal entropies. A simple, but non-unique way of fixing this problem is to use normalized version of the metric κ. There are two ways of normalizing the metric:

$$\nu(X, Y, Z) \stackrel{\text{def}}{=} \frac{\rho(X, Y) + \rho(Y, Z) + \rho(X, Z)}{2H(X, Y, Z)} \quad (19)$$

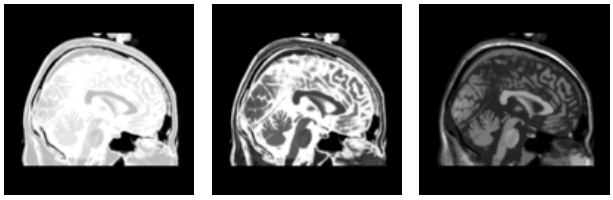
and

$$\sigma(X, Y, Z) \stackrel{\text{def}}{=} \frac{\rho(X, Y) + \rho(Y, Z) + \rho(X, Z)}{2(H(X) + H(Y) + H(Z))}, \quad (20)$$

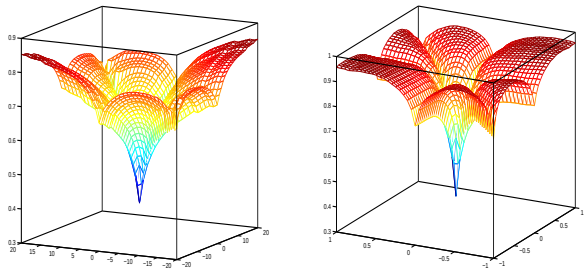
where it can be shown that $0 \leq \sigma(X,Y,Z) \leq 1$. If $X = Y = Z$ then $\sigma(X,Y,Z) = 0$ and if X, Y and Z are independent of each other, $\sigma(X,Y,Z) = 1$. These also hold for $\nu(X,Y,Z)$.

Since we do not want to estimate the joint probability mass function $p(X,Y,Z)$ of three random variables X, Y and Z, we use (20) instead of (19) in our experiments.

In Figure 4, we plot the values of the measure σ for a PD image, a rotated MR T2 and a rotated MR T1 image, where both rotation angles range from -20 degrees to 20 degrees and the values of the metric σ for a PD image, scaled MR T2 and scaled MR T1 image, where both scale ranges are from -1 to 1. From Figure 4, the normalized measure σ achieves its minimum at the point where rotation or scale is zero. Once again this anecdotally demonstrates that we can minimize σ to recover transformations of rotation or scale.



(a) PD image (b) MR T2 image (c) MR T1 image



(d) σ with rotation (e) σ with scaling

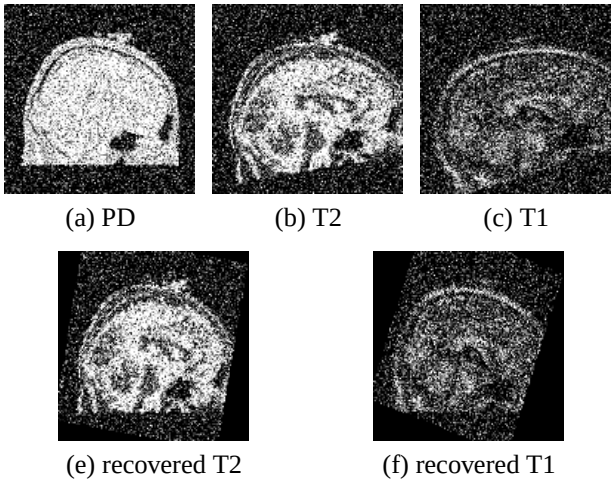
Figure 4: σ with rotation and scaling of PD, rotated T2 and rotated T1 images

3 Experimental Results

3.1 Affine Registration of PD, T2 and T1 MR 2D images

In all our medical imaging experiments, we use the powerful Brainweb simulated MRI volumes for a normal brain [10]. The simulator is based on an anatomical model of a normal brain. The main advantage of using simulated MR data is that the ground truth is known.

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.



(a) PD (b) T2 (c) T1

(e) recovered T2 (f) recovered T1

Figure 5: 2D slice along sagittal direction

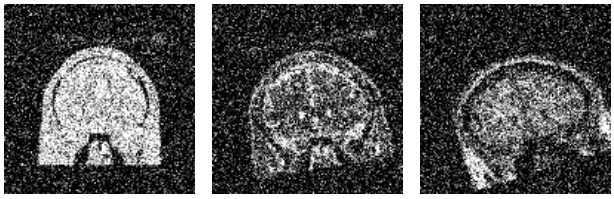
	T on MR T2		T on MR T1	
	ground truth	results	ground truth	results
s	0.1	0.1	0.2	0.18
t	0.1	0.1	0.2	0.18
θ	5	5	10	9.4
ϕ	5	5	10	9.8
e	3	3	5	5
f	3	3	5	5

Table 1: results of 2D slice along sagittal direction

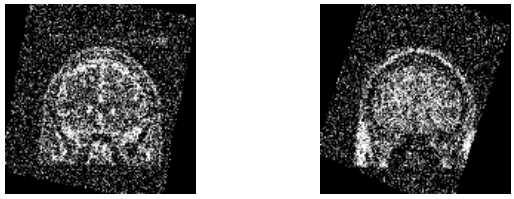
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2^s & 0 \\ 0 & 2^s \end{bmatrix} R(\theta) \begin{bmatrix} 2^t & 0 \\ 0 & 2^{-t} \end{bmatrix} R(\phi)$$

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] and the range of translations is [-10 10].

In the experiments we use 3 triplets of 2D slices of 3D PD, T2 and T1 MR brain volume images. The slices are chosen in sagittal, coronal and axial directions. We then transform the T2 image with an affine transformation $\hat{T}_1$. The scale and shear parameters are 0.1 and 0.1 respectively, and the two rotation parameters are 5 and 5 degrees respectively, with the two translations along the x and y direction being 3 and 3 pixels respectively. Hence



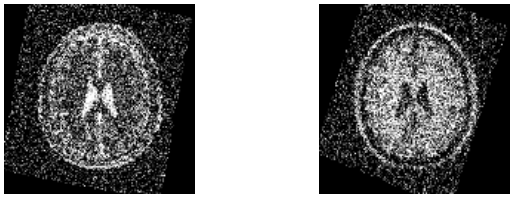
(a) PD (b) T2 (c) T1



(e) recovered T2 (f) recovered T1

Figure 6: 2D slice along coronal direction

(a) PD (b) T2 (c) T1



(e) recovered T2 (f) recovered T1

Figure 7: 2D slice along axial direction

	T on MR T2		T on MR T1	
	ground truth	results	ground truth	results
s	0.1	0.1	0.2	0.16
t	0.1	0.1	0.2	0.22
θ	5	5.4	10	10.6
ϕ	5	5.4	10	10.4
e	3	3	5	5
f	3	3	5	5

Table 2: results of 2D slice along coronal direction

	T on MR T2		T on MR T1	
	ground truth	results	ground truth	results
s	0.1	0.12	0.2	0.18
t	0.1	0.1	0.2	0.2
θ	5	5	10	9.4
ϕ	5	5.4	10	10.2
e	3	3	5	5
f	3	3	5	5

Table 3: results of 2D slice along axial direction

$\hat{T}_1 = \begin{bmatrix} 1.1324 & -0.1866 & 0 \\ 0.1866 & 0.9837 & 0 \\ 3 & 3 & 1 \end{bmatrix}$, with $s_1 = 0.1$, $t_1 = 0.1$, $\theta_1 = 5$, $\phi_1 = 5$, $e_1 = 3$ and $f_1 = 3$. We then transform the T1 MR image with an affine transformation $\hat{T}_2$. The scale and shear parameters are 0.3 and -0.1 respectively and the two rotation parameters are 10 and -5 degrees respectively, with the translations along the x and y directions being 5 and -3 pixels respectively. Hence $\hat{T}_2 = \begin{bmatrix} 1.2496 & -0.3967 & 0 \\ 0.3967 & 0.9301 & 0 \\ 5 & 5 & 1 \end{bmatrix}$, with $s_2 = 0.2$, $t_2 = 0.2$, $\theta_2 = 10$, $\phi_2 = 10$, $e_2 = 5$ and $f_2 = 5$. We add Gaussian noise with zero mean and standard deviation 0.1 to the PD, transformed T2 and transformed T1 MR images and register the three images simultaneously. (The original intensity range of all images is normalized to the [0 1] interval). We use a coarse-to-fine search strategy to find the optimal T_1^* and T_2^*. The registration measure σ is computed only in the overlap area of the three images with nearest neighbor interpolation used for transforming the image intensities.

In all results shown here, the normalized measure σ was used. Figure 5 shows 2D PD, transformed T2 and transformed T1 slices along the sagittal direction and the registration results. Table 1 shows the registration results of these sagittal slices. Figure 6 shows 2D PD, transformed T2 and transformed T1 slices along coronal direction and registration results. Table 2 shows the registration results of these coronal slices. Figure 7 shows 2D PD, transformed T2 and transformed T1 slices along axial direction and registration results. Table 3 shows the registration results of these axial slices.

These experiments above demonstrate a proof of concept of our approach to multimodality image registration since T_1 and T_2 are being simultaneously determined. Future work will focus on validation studies from which we hope to elicit capture range and tolerance to noise, etc.

3.2 Matching face images obtained under different illuminations

In the following experiments, we use face images from the AR Face Database [11]. We simultaneously register three face images with different illuminations. Essentially, either one side of the face or the other or none

(a) left light on (b) right light on (c) no light on

(e) recovered right light on (f) recovered no light on

Figure 8: Images under different lighting conditions (left light on, right light on, and no light on) and wearing sun glasses

	T on right light on		T on no light on	
	ground truth	results	ground truth	results
s	-0.2	-0.22	0.2	0.14
t	-0.2	-0.2	0.2	0.14
θ	-10	-9.6	10	9.6
ϕ	-10	-10	10	10.6
e	-5	-5	5	5
f	-5	-5	5	5

Table 4: results under different lighting conditions (left light on, right light on, and no light on) and wearing sun glasses

are exposed to a light source. We use affine transformations $\hat{T}_1 = \begin{bmatrix} 0.7049 & 0.3006 & 0 \\ -0.3006 & 0.9740 & 0 \\ -5 & -5 & 1 \end{bmatrix}$ and $\hat{T}_2 = \begin{bmatrix} 1.2496 & -0.3967 & 0 \\ 0.3967 & 0.9301 & 0 \\ 5 & 5 & 1 \end{bmatrix}$ to transform two of the face images. Registration is performed against the one untransformed image. The true affine transformation parameters are $s_1 = -0.2$, $t_1 = -0.2$, $\theta_1 = -10$, $\phi_1 = -10$, $e_1 = -5$, $f_1 = -5$, $s_2 = 0.2$, $t_2 = 0.2$, $\theta_2 = 10$, $\phi_2 = 10$, $e_2 = 5$ and $f_2 = 5$.

In all subsequent results, the normalized measure σ was used. Figure 8 shows left-light-on, transformed right-light-on and no-light-on face images wearing sun glasses and the corresponding registration results recovered using the normalized measure σ. Table 4 shows the registration results of these face images wearing sun glasses under different light conditions. Figure 9 shows left-light-on, transformed right-light-on and no-light-on face images wearing a scarf and the registration results. Table 6 shows the registration results of these face images wearing a scarf under different light conditions. Figure 10 shows left-light-on, transformed right-light-on and no-light-on face images wearing sun glasses or a scarf and results recovered using the normalized measure σ. Table 6 shows the registration results of these face images wearing sun glasses or a scarf under different light conditions.

(a) left light on (b) right light on (c) no light on

(e) recovered right light on (f) recovered no light on

Figure 9: Images under different lighting conditions (left light on, right light on, and no light on) and wearing a scarf

	T on right light on		T on no light on	
	ground truth	results	ground truth	results
s	-0.2	-0.20	0.2	0.16
t	-0.2	-0.22	0.2	0.2
θ	-10	-10.6	10	10.2
ϕ	-10	-9.6	10	10.2
e	-5	-5	5	5
f	-5	-5	5	5

Table 5: results under different lighting conditions (left light on, right light on, and no light on) and wearing a scarf

Once again, these three experiments demonstrate a proof of concept with more validation experiments required to better understand the performance under different illuminations.

4 Conclusions

We have presented an information metric for multi-modality image registration, which works very well for affine registration. And we have clearly demonstrated that this metric can be easily extended to the multiple image 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. Intuitively, in the case of two image registration, the metric measures the amount of ignorance in one image given the other and vice versa. The transformation that reduces this "ignorance" metric is deemed optimal. In the multiple image

(a) left light on (b) right light on (c) no light on

(e) recovered right light on (f) recovered no light on

Figure 10: Images under different lighting conditions and wearing sun glasses or scarf

	T on right light on		T on no light on	
	ground truth	results	ground truth	results
s	-0.2	-0.24	0.2	0.14
t	-0.2	-0.2	0.2	0.14
θ	-10	-9.4	10	9.6
ϕ	-10	-9.6	10	10.6
e	-5	-5	5	5
f	-5	-5	5	5

Table 6: results under different lighting conditions (left light on, right light on, and no light on) and wearing sun glasses or a scarf

case, the metric measures the amount of ignorance in one image given the other images. Since it is computationally expensive to compute the conditional entropy (relative ignorance) of one image given the rest, we instead use an upper bound of this entropy. The upper bound is also a metric and is not computationally expensive. In the case of registering two images, the computational expense of our metric is the same as using mutual information. In the case of registering three images, since an upper bound is used, we do not need to estimate the joint probability mass function of three image intensities. The upper bound requires us to estimate 3 joint probability mass functions of two image intensities. We stress that a similar extension of mutual information to three images is not straightforward. In our future work, we plan to extend the use of this metric to nonrigid multimodality registration. We also plan to use an efficient density estimator to estimate high dimensional density functions so that we can directly use the metric instead of merely using an upper bound.

Acknowledgments

This work is supported by NSF grants 0196457 and 0307712. We thank Tim Cootes for pointing out the difficulty of extending mutual information from the intermodality to the multimodality case.

References

[1] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Fifth Intl. Conf. Computer Vision (ICCV)*. 1995, pp. 16–23, IEEE Press.

[2] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3D multi–modality medical image registration using feature space clustering," in *Computer Vision, Virtual Reality and Robotics in Medicine*, N. Ayache, Ed. 1995, vol. 905 of *Lecture Notes in Computer Science*, Springer–Verlag.

[3] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[4] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, NY, 1991.

[5] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Incorporating connected region labelling into automated image registration using mutual information," in *Mathematical Methods in Biomedical Image Analysis*, A. A. Amini, F. L. Bookstein, and D. C. Wilson, Eds., pp. 23–31. IEEE Computer Soc. Press, 1996.

[6] J. L. Boes and C. R. Meyer, "Multi-variate mutual information for registration," in *Medical Image Computing and Computer-Assisted Intervension*, C. Taylor and A. Colchester, Eds., vol. 1679 of *Lecture notes in Computer Science*, pp. 606–612. Springer-Verlag, 1999.

[7] J. A. Lynch, C. G. Peterfy, D. L. White, R. A. Hawkins, and H. K. Genant, "MRI-SPECT image registration using multiple MR pulse sequences to examine osteoarthritis of the knee," in *Medical Imaging: Image Processing*, K. M. Hanson, Ed., vol. 3661 of *Proc. SPIE*, pp. 68–77. SPIE, 1999.

[8] C. Rajski, "A metric space of discrete probability distributions," *Information and Control*, vol. 4, pp. 371–377, 1961.

[9] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognit.*, vol. 32, no. 1, pp. 71–86, 1999.

[10] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, 1998.

[11] A. M. Martinez and R. Benavente, "The AR Face Database," CVC Technical Report No.24, 1998.