

Normalized similarity measures for medical image registration

A. Bardera, M. Feixas and I. Boada

Institut d'Informàtica i Aplicacions, Universitat de Girona,
Campus Montilivi, 17071, Girona, Spain

ABSTRACT

Two new similarity measures for rigid image registration, based on the normalization of Jensen's difference applied to Rényi and Tsallis-Havrda-Charvát entropies, are introduced. One measure is normalized by the first term of Jensen's difference, which in our proposal coincides with the marginal entropy, and the other by the joint entropy. These measures can be seen as an extension of two measures successfully applied in medical image registration: the mutual information and the normalized mutual information. Experiments with various registration modalities show that the new similarity measures are more robust than the normalized mutual information for some modalities and a determined range of the entropy parameter. Also, a certain improvement on accuracy can be obtained for a different range of this parameter.

Keywords: Multi-modal registration, generalized entropies, Jensen's difference, information theory, mutual information, similarity measures

1. INTRODUCTION

Multimodal image registration is essential in medical image processing. The definition of accurate and robust similarity measures, which enable us to evaluate the spatial correspondence between images, has become a challenging research area. Among the different image registration measures proposed, mutual information (MI)¹ and normalized mutual information (NMI)² are the most commonly used due to their accuracy, robustness, and universality. MI measures the statistical dependence between images and NMI improves the robustness of MI avoiding some misregistrations.

In this paper, we introduce two new normalized similarity measures based on the Jensen difference,^{3,4} also called generalized mutual information (GMI), applied to Rényi and Tsallis-Havrda-Charvát entropies. Both measures are a normalization of GMI and can be seen as an extension of MI and NMI . The first measure, $nGMI$, is normalized by the first term of the Jensen difference, which in our proposal is the marginal entropy, and the second, $NGMI$, by the joint entropy. It is important to remark that one particular case of $NGMI$ is NMI . The behavior of our measures is studied in rigid intrasubject registration of head images acquired using tomographic modalities. The results obtained with various registration modalities (MR to synthetic MR, MR to CT, MR to PET) show that the new similarity measures are more robust than the normalized mutual information for some modalities and determined parameters of the entropies. We have also to emphasize that, for a parameter $\alpha = 2$, the Jensen difference using the Tsallis-Havrda-Charvát entropy shows an acceptable performance and is less demanding than using the Shannon entropy.

This paper is organized as follows: Section 2 presents the most fundamental information theoretic definitions used in medical image registration and the Jensen difference applied to Rényi and Tsallis-Havrda-Charvát entropies. Two new normalized similarity measures based on the Jensen difference are introduced in Section 3 and their performance is evaluated in Section 4. Finally, in Section 5 we present our conclusions and future work.

Further author information: (Send correspondence to A.B.)
E-mail: {anton.bardera,feixas,imma}@ima.udg.es

©Copyright 2004 Society of Photo-Optical Instrumentation Engineers
This paper was published in Medical Imaging SPIE 2004 and is made available as an electronic reprint with the permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or others means, duplication of any material in this paper for a fee or commercial purposes, or modification of the content of the paper are prohibited.

2. INFORMATION THEORETIC TOOLS

In 1948, Claude Shannon published “A mathematical theory of communication”⁵ which marks the beginning of information theory. In this paper, he defined measures such as entropy and mutual information, and introduced the fundamental laws of data compression and transmission. Information theory deals with the processing of *information* and is used in fields such as physics, computer science, statistics, biology, linguistics, neurology, learning, etc..⁶ Over the last decade, information theoretic measures have been successfully applied to medical image registration.⁷⁻⁹

In this section we review the most basic definitions of information theory⁶ and their application to medical image registration. We also introduce other essential concepts for our approach: the Jensen inequality and its application to the Rényi and Tsallis-Havrda-Charvát entropies.

2.1. Shannon Entropy and Mutual Information

The *Shannon entropy* $H(X)$ of a discrete random variable X with values in the set $S = \{x_1, x_2, \dots, x_n\}$ is defined as

$$H(X) = - \sum_{i=1}^n p_i \log p_i, \quad (1)$$

where $n = |S|$, $p_i = Pr[X = x_i]$ for $i \in \{1, \dots, n\}$. The logarithms are taken in base 2 and entropy is expressed in bits. The convention that $0 \log 0 = 0$ is used. We can use interchangeably the notation $H(X)$ or $H(p)$ for the entropy, where $p = \{p_1, p_2, \dots, p_n\}$ is the corresponding probability distribution. As $-\log p_i$ represents the information associated with the result x_i , the entropy gives us the average information or uncertainty of a random variable.

If we consider another random variable Y with marginal probability distribution q , corresponding to values in the set $S' = \{y_1, y_2, \dots, y_m\}$, the *joint entropy* of a pair of discrete random variables (X, Y) is defined as

$$H(X, Y) = - \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log p_{ij}, \quad (2)$$

where $m = |S'|$ and $p_{ij} = Pr[X = x_i, Y = y_j]$ is the joint probability, and the *conditional entropy* is defined as

$$H(X|Y) = - \sum_{j=1}^m \sum_{i=1}^n p_{ij} \log p_{i|j}, \quad (3)$$

where $p_{i|j} = Pr[X = x_i | Y = y_j]$ is the conditional probability. The Bayes theorem, $p_{ij} = p_i p_{j|i} = q_j p_{i|j}$, expresses the relation between the different kind of probabilities. If X and Y are independent, then $p_{ij} = p_i q_j$. Conditional entropy can be thought of in terms of an information channel whose input is the random variable X and whose output is the random variable Y . $H(X|Y)$ corresponds to the uncertainty in the channel input from the receiver’s point of view, and vice versa for $H(Y|X)$. Note that in general $H(X|Y) \neq H(Y|X)$.

The *mutual information* between two random variables X and Y is defined as

$$I(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = \sum_{i=1}^n \sum_{j=1}^m p_{ij} \log \frac{p_{ij}}{p_i q_j}, \quad (4)$$

Mutual information represents the amount of information that one random variable, the output of the channel, contains about a second random variable, the input of the channel, and vice versa. $I(X, Y)$ is a measure of the shared information or dependence between X and Y .

The following properties are met:

1. $H(X) \geq H(X|Y) \geq 0$
2. $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$
3. $I(X, Y) = H(X) + H(Y) - H(X, Y) \geq 0$ with equality if, and only if, X and Y are independent.

2.2. Mutual Information for Image Registration

Different approaches for automatic image registration have been proposed in medical imaging.⁷ The most successful methods are based on mutual information, which is a measure of the dependence between two images.

The registration of two images is represented by an information channel $X \rightarrow Y$, where the random variables X and Y represent, respectively, the images A and B . Their marginal probability distributions, $\{p_i\}$ and $\{q_j\}$, and the joint probability distribution, $\{p_{ij}\}$, are obtained by simple normalization of the marginal and joint intensity histograms of the overlapping areas of both images. The conditional probability can be calculated using the Bayes theorem, leading to the transition probability matrix P of the channel (conditional probability matrix):

$$P = \begin{pmatrix} p_{1|1} & p_{2|1} & \cdots & p_{m|1} \\ p_{1|2} & p_{2|2} & \cdots & p_{m|2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1|n} & p_{2|n} & \cdots & p_{m|n} \end{pmatrix}, \quad (5)$$

where n and m are, respectively, the number of bins of the intensity histograms of images A and B . The row i of this matrix is represented by

$$P_i = (p_{1|i} \quad p_{2|i} \quad \cdots \quad p_{m|i}). \quad (6)$$

For the inverse channel $Y \rightarrow X$, Q and Q_j are, respectively, the conditional probability matrix and the row j of this matrix.

From these probability distributions, the joint entropy $H(X, Y)$, which measures the dispersion of the joint histogram, and the mutual information $I(X, Y)$, which quantifies how well an image explains another, can be calculated. A registration criterion based on the minimization of joint entropy was proposed by Collignon et al.¹⁰ and Studholme et al.¹¹ The main drawback of this method is its high sensitivity to the overlap area. The registration method based on the maximization of mutual information, almost simultaneously introduced by Maes and Collignon et al.¹² and Viola and Wells et al.,¹ solves partially the above mentioned problem. This technique is based on the conjecture that the correct registration corresponds to the maximum mutual information between the overlapping areas of the two images. Later, Studholme et al.² proposed a normalized measure of mutual information defined by

$$NMI(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)} = 1 + \frac{I(X, Y)}{H(X, Y)}, \quad (7)$$

which is more robust than MI , due to its greater independence of the overlap area. These two measures are widely used in numerous studies and papers (see surveys^{8,9}).

2.3. Jensen's Difference

In this section, Jensen's inequality, generalized entropies and generalized mutual information are briefly introduced.

2.3.1. Jensen's Inequality

A function $f(x)$ is *convex* over an interval $[a, b]$ (the graph of the function lies below any chord) if for every $x_1, x_2 \in [a, b]$ and $0 \leq \lambda \leq 1$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2). \quad (8)$$

A function is strictly convex if equality holds only if $\lambda = 0$ or $\lambda = 1$. A function $f(x)$ is *concave* (the graph of the function lies above any chord) if $-f(x)$ is convex. For instance, x^2 and $x \log x$ (for $x \geq 0$) are strictly convex functions, and $\log x$ (for $x \geq 0$) is a strictly concave function.⁶

A generalization of the above convexity property, called *Jensen's inequality*, is widely used in mathematics, information theory, and different engineering areas as a divergence measure. If f is a convex function on the interval $[a, b]$, then

$$\sum_{i=1}^n \lambda_i f(x_i) - f\left(\sum_{i=1}^n \lambda_i x_i\right) \geq 0, \quad (9)$$

where $0 \leq \lambda \leq 1$, $\sum_{i=1}^n \lambda_i = 1$, and $x_i \in [a, b]$. If f is a concave function, the inequality is reversed.

2.3.2. Generalized Entropies

Two different generalizations of the Shannon entropy (Rényi and Tsallis-Havrda-Charvát entropies) have been proposed.^{3,4} Both are used in image registration.^{13,14}

The *Rényi entropy* is defined by

$$H_\alpha^R(X) = \frac{1}{1-\alpha} \log \sum_{i=1}^n p_i^\alpha, \quad (10)$$

where $\alpha > 0$ and $\alpha \neq 1$. When $\alpha \rightarrow 1$, $H_\alpha^R(X) = H(X)$. $H_\alpha^R(X)$ is a concave function of p if $\alpha \leq 1$, but neither concave nor convex if $\alpha > 1$.

The *Tsallis-Havrda-Charvát entropy* is defined by

$$H_\alpha^T(X) = \frac{1}{\alpha-1} \left(1 - \sum_{i=1}^n p_i^\alpha\right), \quad (11)$$

where $\alpha > 0$ and $\alpha \neq 1$. When $\alpha \rightarrow 1$, $H_\alpha^T(X) = H(X)$. $H_\alpha^T(X)$ is a concave function of p for $\alpha > 0$. When $\alpha = 2$, $H_2^T(X)$ becomes the *Gini-Simpson index* of diversity,

$$H_2^T(X) = 1 - \sum_{i=1}^n p_i^2, \quad (12)$$

which has been widely used in learning.¹⁵

2.3.3. Generalized Mutual Information

From Jensen's inequality (9), the concavity of the entropy functions provides us with the following inequality³ :

$$J_\alpha^h(\{\Pi_i\} : \{\pi_i\}) = H_\alpha^h\left(\sum_{i=1}^n \pi_i \Pi_i\right) - \sum_{i=1}^n \pi_i H_\alpha^h(\Pi_i) \geq 0, \quad (13)$$

where $\Pi_1, \Pi_2, \dots, \Pi_n$ are a set of probability distributions and $\pi_1, \pi_2, \dots, \pi_n$ are the priori probabilities or weights, fulfilling $\sum_{i=1}^n \pi_i = 1$. The superindex h can be substituted by R (Rényi entropy) or T (Tsallis-Havrda-Charvát entropy), depending on the entropy considered. $J_\alpha^h(\{\Pi_i\} : \{\pi_i\})$ is called Jensen's difference and is an important divergence measure between the set of distributions $\{\Pi_i\}$. For the Rényi entropy, the parameter α can only take values in the range $0 < \alpha \leq 1$ due to the fact that $H_\alpha^R(X)$ is neither concave nor convex if $\alpha > 1$. The particular case $\pi_i = \frac{1}{n}$ for the Jensen difference applied to the Rényi entropy has been studied by He et al.¹⁴ in medical image registration.

Jensen's difference coincides with $I(X, Y)$ when $\alpha = 1$, $\{\pi_i\}$ is the marginal probability distribution $\{p_i\}$ of X , and $\{\Pi_i\}$ are the rows $\{P_i\}$ of the probability conditional matrix of the channel (5). Hence, $J_\alpha^h(\{\Pi_i\} : \{\pi_i\})$ is called *generalized mutual information*.³ In this paper, it is also denoted by *GMI*.

3. NORMALIZED SIMILARITY MEASURES

The goal of the registration of two images A and B is to obtain the transformation T that maximizes the similarity between A and $T(B)$. Thus, one of the cornerstones of the image registration is the measure that quantifies the similarity. The similarity measures we present in this paper are based on Jensen's difference (13).

As we have seen in Sect.2, Jensen's difference coincides with $I(X, Y)$ when $\alpha = 1$, $\{\pi_i\}$ is the marginal probability distribution $\{p_i\}$ of X , and $\{\Pi_i\}$ are the rows $\{P_i\}$ of the probability conditional matrix of the channel. In the following proposed measures, the conditions $\pi_i = p_i$ and $\Pi_i = P_i$ are fixed. Thus, only two parameters, α and h , can be changed. From these assumptions, for the channel $X \rightarrow Y$, the generalized mutual information is expressed by

$$\begin{aligned} GMI_\alpha^h(X \rightarrow Y) &= J_\alpha^h(\{P_i\} : \{p_i\}) = H_\alpha^h\left(\sum_{i=1}^n p_i P_i\right) - \sum_{i=1}^n p_i H_\alpha^h(P_i) \\ &= H_\alpha^h(q) - \sum_{i=1}^n p_i H_\alpha^h(P_i) = H_\alpha^h(Y) - \sum_{i=1}^n p_i H_\alpha^h(P_i). \end{aligned} \quad (14)$$

For the reverse channel $Y \rightarrow X$, we have

$$GMI_\alpha^h(Y \rightarrow X) = J_\alpha^h(\{Q_j\} : \{q_j\}) = H_\alpha^h(X) - \sum_{j=1}^m q_j H_\alpha^h(Q_j). \quad (15)$$

In general, this measure is not symmetric with respect to the reversion of the channel. Thus, $GMI_\alpha^h(X \rightarrow Y) \neq GMI_\alpha^h(Y \rightarrow X)$.

Although a simple substitution of MI for GMI_α^h can be used as an absolute similarity measure between two images, in this paper we focus our interest on a relative one. Such a decision is motivated by the better behavior of NMI with respect to MI .²

Next, two new similarity measures for rigid image registration, $nGMI$ and $NGMI$, based on the normalization of GMI_α^h , are proposed. In order to register the images, the transformation which maximizes the $nGMI$ or $NGMI$ has to be found.

3.1. nGMI

The first natural choice is to normalize $GMI_\alpha^h(X \rightarrow Y)$ or $GMI_\alpha^h(Y \rightarrow X)$ by its corresponding first term in formulae (14) and (15), i.e., the marginal probability distribution. When $\alpha = 1$, this ratio is given by $I(X, Y)/H(Y)$ or $I(X, Y)/H(X)$. This is a correlation measure⁶ between the random variables X and Y . From the two possibilities, we take the value which gives us the maximum correlation between the images. So, we postulate that the best choice is always taking the maximum possible correlation. Clearly, we obtain a measure independent of the order in which the images are taken in the registration process.

From these assumptions, the normalization of GMI_α^h by the marginal distribution is given by

$$nGMI_\alpha^h = \max \left\{ \frac{GMI_\alpha^h(X \rightarrow Y)}{H_\alpha^h(Y)}, \frac{GMI_\alpha^h(Y \rightarrow X)}{H_\alpha^h(X)} \right\}. \quad (16)$$

This measure can be called *generalized correlation*.

The most relevant properties of this measure are:

- It takes values in $[0, 1]$.
- It is symmetric, i.e., the result of the registration is independent of the image selected as reference.
- If $\alpha = 1$, then $nGMI = \frac{I(X, Y)}{\min\{H(X), H(Y)\}}$. A similar measure, though different, based on the Kolmogorov complexity, has been applied by Li et al.¹⁶

3.2. NGMI

Similarly to Studholme’s approach, GMI_α^h can be normalized by the corresponding joint entropy $H_\alpha^h(X, Y)$. Studholme’s normalization is well justified by Butz and Thiran¹⁷ from a different perspective. Another justification about its good behavior can be given. In information theory, the most basic information metric is expressed by $H(X, Y) - I(X, Y) = H(X|Y) + H(Y|X)$. This measure fulfils the requirements for a true distance.⁶ Thus, $\frac{H(X, Y) - I(X, Y)}{H(X, Y)} = 1 - \frac{I(X, Y)}{H(X, Y)}$ is a normalized metric and NMI can be considered as a normalized similarity metric. In image registration literature, this fact has not been sufficiently stressed.

The normalization of GMI_α^h by the joint entropy is given by

$$NGMI_\alpha^h = \frac{\max\{GMI_\alpha^h(X \rightarrow Y), GMI_\alpha^h(Y \rightarrow X)\}}{H_\alpha^h(X, Y)}. \quad (17)$$

The most relevant properties of this measure are:

- It takes values in $[0, 1]$.
- It is symmetric.
- If $\alpha = 1$, then $NGMI = \frac{I(X, Y)}{H(X, Y)}$, i.e., we obtain the NMI measure.

A similar approach was taken by Wachowiak et al,¹³ where the mutual information of Rényi and Tsallis-Havrda-Charvát is normalized by the corresponding joint entropy. It is important to remark that the GMI_α^h does not coincide with the Rényi and Tsallis-Havrda-Charvát mutual information. Thus, our $NGMI_\alpha^h$ criterion is different from the one introduced by them.

4. RESULTS

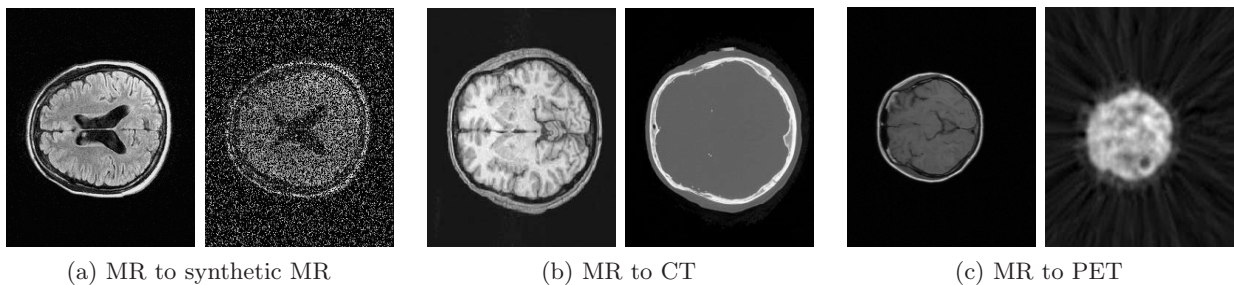


Figure 1. Test image pairs.

The performance of the $nGMI$ and $NGMI$ criteria has been tested for rigid intrasubject registration of head images. The robustness of these measures have been evaluated in terms of the partial image overlap. This has been done using the parameter AFA (Area of Function Attraction) introduced by Capek et al.¹⁸ This parameter evaluates the range of convergence of a registration measure to its global maximum, counting the number of pixels, i.e. x-y translations in image space, from which the global maximum is reached by applying a maximum gradient method. Note that this global maximum may not necessarily be the optimal registration position. The AFA parameter represents the robustness with respect to the different initial positions of the images to be registered and with respect to the convergence to a local maximum of the similarity measure that leads to an incorrect registration. The higher the AFA, the wider the attraction basin of the measure is.

Three different sets of experiments in 2-D registration, corresponding to three modalities (MR to synthetic MR, MR to CT and MR to PET), have been evaluated. The images used in our study have been obtained from the Vanderbilt data base and from the Josep Trueta Hospital in Girona. The resolution of the MR and CT images is 256×256 and the PET images is 128×128 . All the experiments have been done using a histogram of

256 bins. A reduction of the number of bins decreases the cost of the algorithm and can eventually increase the robustness of the similarity measures. In this paper we do not analyze the behavior of our measures depending on the segmentation of the histogram, nor do we study the dependence of our measures on different interpolation schemes and speed-up strategies. As the standard algorithms are implemented to find the global minimum, in our experiments, all measures have been multiplied by -1 to have a minimum in the registration position. Next we present our results.

4.1. MR to noisy MR

The first experiment tests the performance of our measures when registering a pair of images composed of a MR and a synthetically generated MR image. This synthetic image is created from the original MR image by swapping pairs of randomly selected pixels (see Fig. 1(a)).

Figure 2 shows the graphs of the parameter AFA for $nGMI$ and $NGMI$ using Rényi and Tsallis entropies and the attraction basin for $NGMI_{1.5}^T$. Note that α values for the Rényi entropy go from 0 to 1 and for the Tsallis entropy from 0 to 2. With the Rényi entropy (see Fig. 2(a)), the proposed measures increase monotonically for $0 < \alpha < 0.5$. For α close to 0.6, $nGMI^R$ reaches a maximum and outperforms NMI . For $\alpha > 0.6$, both measures match NMI . In Fig. 2(b), we show the results using the Tsallis entropy. For $\alpha < 1$, $nGMI^T$ and $NGMI^T$ increase monotonically, almost in the same way. For $\alpha = 1$, these measures match NMI and for $\alpha > 1$ they clearly outperform NMI . Special attention must be paid to $\alpha = 2$, because of its low computational cost and good performance.

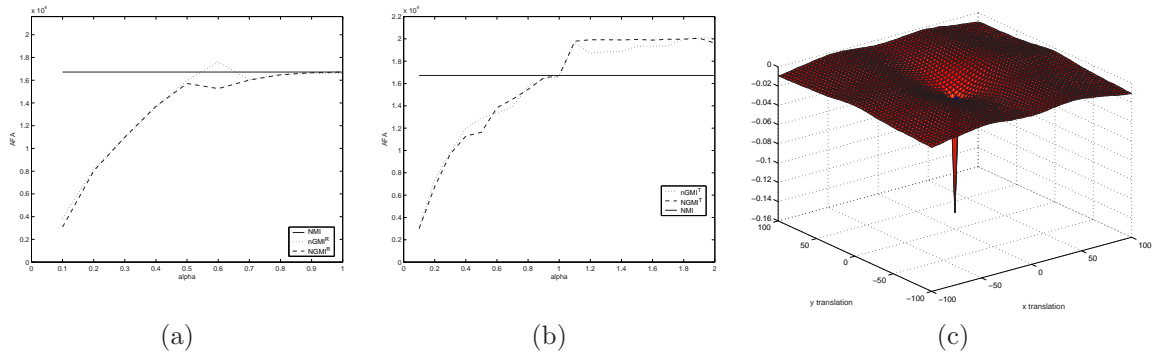


Figure 2. MR to synthetic MR registration. Curves of parameter AFA for (a) $nGMI^R$ and $NGMI^R$ vs NMI and (b) $nGMI^T$ and $NGMI^T$ vs NMI . (c) Attraction basin for $NGMI_{1.5}^T$

In Fig. 3, we illustrate the behavior of the $nGMI^R$ and $NGMI^T$ over various x-shifts and for different α values. For $nGMI^R$, $\alpha = 1$ is the best choice. For $NGMI^T$, $\alpha > 1$ leads to sharper peaks.

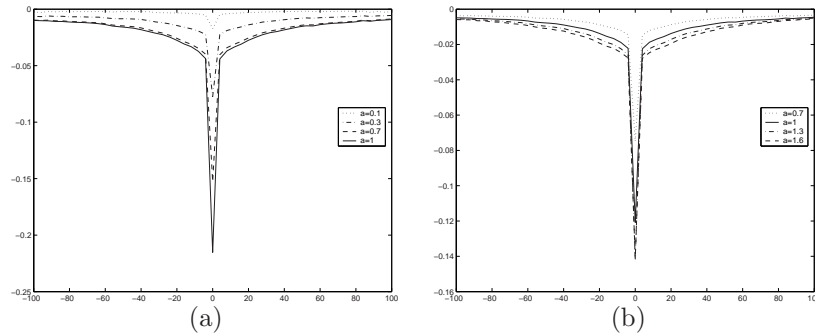


Figure 3. MR to synthetic MR registration. Curves of (a) $nGMI^R$ and (b) $NGMI^T$ over x-shifts and for different α values.

4.2. MR to CT

Next, we consider the problem of registering pairs of CT and MR images. This case is of special interest since it is routinely used in clinical practice. Our testing set was composed of eight different pairs of images. Results for all the image pairs found a similar performance to the case presented in this section which corresponds to the image pair of Fig. 1(b) and Fig. 4.

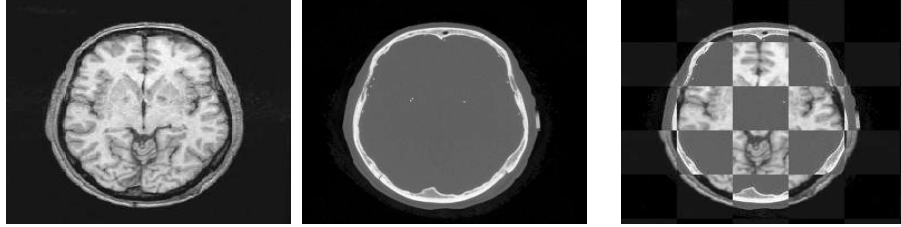


Figure 4. From left to right: MR and CT slices and MR-CT fusion.

Plots of the parameter AFA for $nGMI$ and $NGMI$ using Rényi and Tsallis entropies are shown in Fig. 5. The first graph (see Fig. 5(a)) corresponds to the $nGMI^R$ and $NGMI^R$, with no improvements with respect to NMI . In Fig. 5(b), $nGMI^T$ and $NGMI^T$ outperform NMI for $1.2 < \alpha < 1.8$. The attraction basin for $nGMI_{1.5}^T$ is shown in Fig. 5(c).

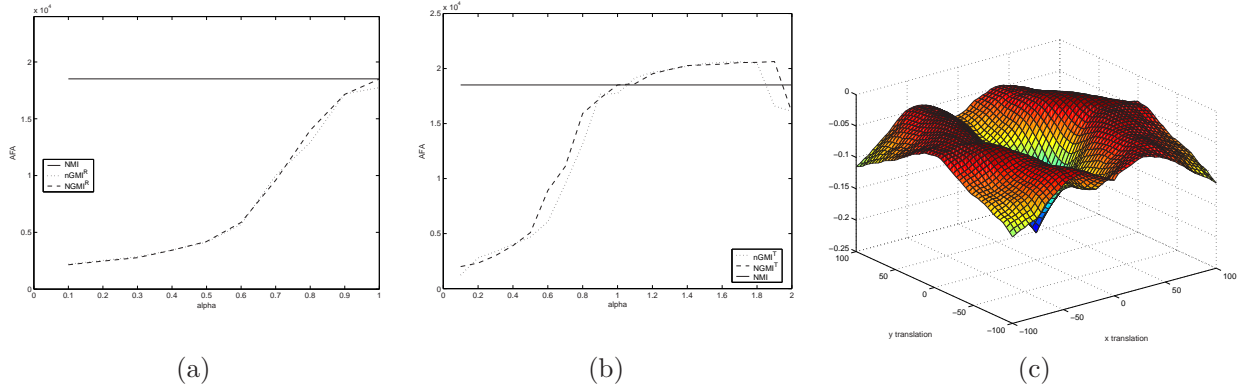


Figure 5. MR to CT registration. Curves of parameter AFA for (a) $nGMI^R$ and $NGMI^R$ vs NMI and (b) $nGMI^T$ and $NGMI^T$ vs NMI . (c) Attraction basin for $nGMI_{1.5}^T$.

Figure 6 illustrates the behavior of our measures using 2-D curves. It can be observed that the attraction basin of the curves increases with α . Small α values lead to narrower registration peaks, and greater α values to smoother curves. Consequently, the best choice is to use high alpha values in the MR to CT registration process.

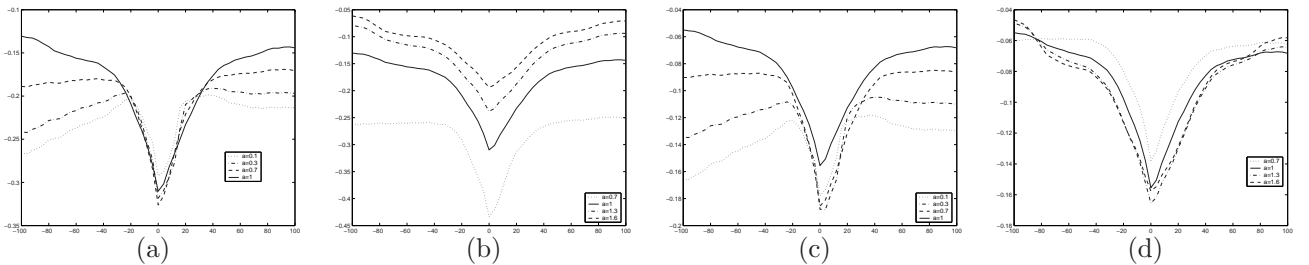


Figure 6. MR to CT registration. Curves of (a) $nGMI^R$ (b) $nGMI^T$ (c) $NGMI^R$ (d) $NGMI^T$ over x-shifts.

4.3. MR to PET

Finally we evaluate the MR to PET multi-modal registration. This is the most difficult case, both because of the low resolution and the large amount of noise of PET images and because of the fewer similarities between MR and PET images. It is quite normal that mutual-information-based matching of these images results in misregistration.

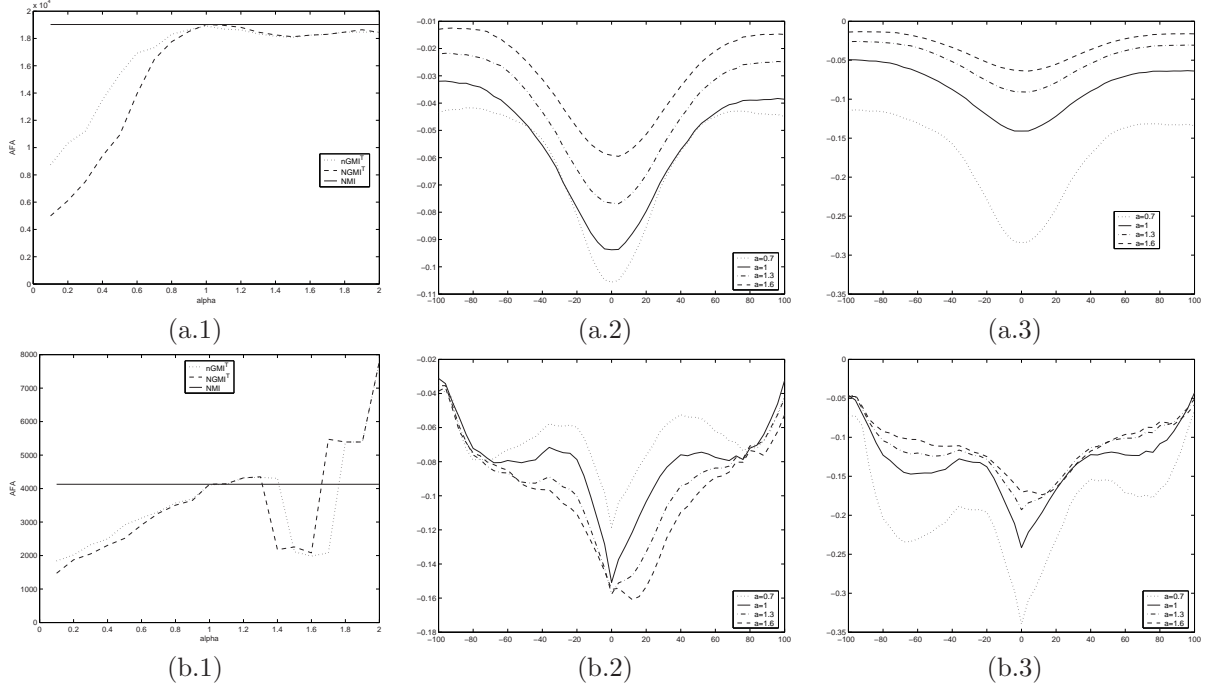


Figure 7. MR to PET registration. Study of two cases (a)(b): (1) AFA for $NGMI^T$ and $nGMI^T$, curves of (2) $NGMI^T$ and (3) $nGMI^T$ over x-shifts.

Different pairs of images have been tested. Figure 7 shows the results obtained for two representative cases. For both cases we plot, from left to right, the parameter AFA for $NGMI^T$ and $nGMI^T$, and also these measures over various x-shifts. We do not show the results obtained using the Rényi entropy because their behavior is similar to the previous tested modalities.

The first row of graphs corresponds to the results obtained with the image pair of Fig. 1(c). The parameter AFA is plotted in Fig. 7.(a.1). Observe that it increases monotonically with α until $\alpha = 1$ and then becomes stable close to NMI . This behavior is similar to the other evaluated registration modalities (see Fig. 2(b) and 5(b)). Analyzing in detail the two next graphs (see Fig. 7.(a.2) and 7.(a.3)), we can observe that, despite the smoothness of $NGMI^T$ and $nGMI^T$ curves, for high α there is a small displacement with respect to the optimal registration point (point 0 in this case). This fact corroborates the ill-defined behavior of these measures for high α values as pointed out by Pluim¹⁹ for NMI . Note that NMI is a particular case of $NGMI$. This effect is more pronounced in the second pair of evaluated images (see Fig. 7.(b.2) and 7.(b.3)). From Fig. 7.(b.1) we see that the AFA for $1.2 < \alpha < 1.7$ decreases until a local minimum is reached and then it increases. This behavior is due to the coexistence of two local minima which leads to a lower attraction basin, but for higher α values, one of the local minima disappears and AFA recovers its normal behavior.

In conclusion, we want to stress that for high values of α although the AFA increases, the global minimum does not coincide with the optimal registration position. On the other hand, for low values of α , the attraction basin decreases, but the peak of the optimal registration is sharper. This behavior can be interpreted as a certain improvement on accuracy for low α values. It is important to observe that in this case local minima can appear.

5. CONCLUSIONS

In this paper, we introduced two new similarity metrics for rigid medical image registration. These metrics are based on the Jensen's difference applied to Rényi and Tsallis-Havrda-Charvát entropies. Both measures are a normalization of Jensen's difference and can be seen as an extension of MI and NMI . The first measure, $nGMI$, is normalized by the marginal entropy and the second, $NGMI$, by the joint entropy. We have evaluated the performance of our measures in rigid intrasubject registration of head images. The results obtained in this paper demonstrate that $nGMI$ and $NGMI$ are more robust than NMI for a determined range of the entropy parameter. Also, it has been found that, for a different range of the entropy parameter, the proposed measures lead to a minimum closer to the optimal registration point than NMI . Depending on the registration modality, we suggest to use one range or the other.

Our future work will be focused on the study of the dependence of our measures on different interpolation schemes and speed-up strategies. Also, we will analyze other possible weights for Jensen's difference. Finally, we will address the registration problem of more than two images using Jensen's difference.

ACKNOWLEDGEMENTS

Some of the images used in our experiments were provided as part of the project, "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 1 R01 NS33926-01, Principal Investigator Prof. J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN. We thank the Institut de Diagnòstic per la Imatge of the Josep Trueta Hospital of Girona for supplying images and its assistance. This project has been funded in part with grant numbers TIC2001-2416-C03-01, TIC2000-1009, TIC2001-2226-C02-02 and 2001-SGR-00296.

REFERENCES

1. P. A. Viola, *Alignment by Maximization of Mutual Information*. PhD thesis, Massachusetts Institute of Technology, Massachusetts (MA), USA, 1995.
2. C. Studholme, *Measures of 3D Medical Image Alignment*. PhD thesis, University of London, London, UK, August 1997.
3. J. Burbea and C. Rao, "On the convexity of some divergence measures based on entropy functions," *IEEE Transactions on Information Theory* **28**(3), pp. 489–495, 1982.
4. C. Rao, *Rao's axiomatization of Diversity Measures (Encyclopedia of Statistical Sciences)*, pp. 614–617. Wiley Interscience, New York (NY), USA, 1983.
5. C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal* **27**, pp. 379–423, 623–656, July, October 1948.
6. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, 1991.
7. D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in Medicine and Biology* **46**, pp. R1–R45, 2001.
8. J. P. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging* **22**, pp. 986–1004, 2003.
9. F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," in *Proceedings of the IEEE*, **91**(10), pp. 1699–1722, 2003.
10. A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal, "3d multi-modality medical image registration using feature space clustering," in *First International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, N. Ayache, ed., *Lecture Notes in Computer Science* **905**, pp. 195–204, Springer-Verlag Vienna-New York, April 1995. Held in Nice, France.
11. C. Studholme, D. Hill, and D. Hawkes, "Multiresolution voxel similarity measures for mr-pet registration," in *Proceedings of the Information Processing in Medical Imaging Conference*, Y. Bizais, ed., pp. 287–298, Kluwer Academic Publishers, (NY), June 1995. Held in Ile de Berder, France.
12. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging* **16**(2), pp. 187–198, 1997.

13. M. P. Wachowiak, R. Smolikova, G. D. Tourassi, and A. S. Elmaghraby, "Similarity metrics based on non-additive entropies for 2d-3d multimodal biomedical image registration," in *Proceedings of SPIE Medical Imaging'03*, pp. 1090–1100, 2003. Held in San Diego, CA.
14. Y. He, A. Ben Hamza, and H. Krim, "A generalized divergence measure for robust image registration," *IEEE Transactions on Signal Processing* **51**, May 2003.
15. T. Elomaa and J. Rousu, "General and efficient multisplitting of numerical attributes," *Machine Learning* **36**, pp. 201–244, 1999.
16. M. Li, X. Chen, X. Li, B. Ma, and P. Vitányi, "The similarity metric," in *Proceedings of the 14th ACM-SIAM Symposium*, pp. 1889–1894, December 2003. Held in Phoenix (AZ), USA.
17. T. Butz, O. Cuisenaire, and J.-P. Thiran, "Multi-modal medical image registration: From information theory to optimization objective," pp. 407–414, July 2002. Proceedings of DSP 2002, Santorini, Greece.
18. M. Capek, L. Mroz, and R. Wegenkittl, "Robust and fast medical registration of 3d-multi-modality data sets," pp. 515–518, June 2001. Proceedings of Medicon 2001, Pula, Croatia.
19. J. P. Plum, *Mutual Information Based Registration of Medical Images*. PhD thesis, Utrecht University, Utrecht, The Netherlands, 2001.