ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Pattern Recognition ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

PATTERN RECOGNITION

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/pr

Multimodality image registration by maximization of quantitative–qualitative measure of mutual information

Hongxia Luan[a], Feihu Qi[a], Zhong Xue[b], Liya Chen[c], Dinggang Shen[b,*]

[a]*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200030, China*
[b]*Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA*
[c]*Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200030, China*

Received 16 August 2006; received in revised form 16 January 2007; accepted 4 April 2007

Abstract

This paper presents a novel image similarity measure, referred to as quantitative–qualitative measure of mutual information (Q-MI), for multimodality image registration. Conventional information measures, e.g., Shannon's entropy and mutual information (MI), reflect quantitative aspects of information because they only consider probabilities of events. In fact, each event has its own utility to the fulfillment of the underlying goal, which can be independent of its probability of occurrence. Thus, it is important to consider both quantitative (i.e., probability) and qualitative (i.e., utility) measures of information in order to fully capture the characteristics of events. Accordingly, in multimodality image registration, Q-MI should be used to integrate the information obtained from both the image intensity distributions and the utilities of voxels in the images. Different voxels can have different utilities, for example, in brain images, two voxels can have the same intensity value, but their utilities can be different, e.g., a white matter (WM) voxel near the cortex can have higher utility than a WM voxel inside a large uniform WM region. In Q-MI, the utility of each voxel in an image can be determined according to the regional saliency value calculated from the scale-space map of this image. Since the voxels with higher utility values (or saliency values) contribute more in measuring Q-MI of the two images, the Q-MI-based registration method is much more robust, compared to conventional MI-based registration methods. Also, the Q-MI-based registration method can provide a smoother registration function with a relatively larger capture range. In this paper, the proposed Q-MI has been validated and applied to the rigid registrations of clinical brain images, such as MR, CT and PET images.

Keywords: Multimodality image registration; Mutual information; Quantitative–qualitative measure of mutual information; Salient measure; Utilities of events

1. Introduction

Registration of multimodality images of the same subject provides a way of fusing different types of information, and it is very important for medical diagnosis and computer-aided surgery [1–14]. For example, by registering MR T1, T2, PD, and FLAIR brain images of the same subject, the white matter lesions (WMLs) in the brain can be identified [1]; by registering the pre- and intra-operative images, computer-aided surgery can be well performed [2]. Mutual information (MI) that measures the statistical dependency between two images has been successfully applied to multimodality image registration [6–9]. Although numerous promising results have been reported, it is worth noting that the MI-based registration methods might have the limited performances, once the initial misalignment of the two images is large or equally the overlay region of the two images is small [10].

Various methods have been proposed to improve the robustness of MI-based registration, and they can be classified into three categories [10]. In the first category, instead of Shannon's entropy, alternative entropies such as Jumarie entropy [12,13] and Rényi entropy [14] have been used. In the second category, normalized MI measures [15,16], which is less sensitive to the changes in the overlap of two images, has been proposed and applied successfully in lots of studies [17–19]. In the third

* Corresponding author.
E-mail addresses: luanhx@sjtu.edu.cn (H. Luan), fhqi@sjtu.edu.cn (F. Qi), zhong.xue@uphs.upenn.edu (Z. Xue), liyachen@sjtu.edu.cn (L. Chen), dinggang.shen@uphs.upenn.edu (D. Shen).

0031-3203/$30.00
doi:10.1016/j.patcog.2007.04.002

category, spatial information has been integrated into the MI-based registration [20–25]. For example, both MI and matching of gradient maps between two images have been used for image registration in Ref. [20]; also, the MI of image features, i.e., gradient, wavelet and other features, is employed for registration [22–25]. Generally, the use of spatial information increases the robustness of registration algorithms.

However, almost all MI-based registration methods treat the voxels of the images equally, when calculating their MI. In fact, different voxels, even having the same intensity [26], should be treated differently since they have different characteristics and utilities on image registration. Salient voxels should have higher utility values, and hence contribute more to determine the transformation between two images. For example, when measuring the MI of two brain images, the white matter (WM) voxels near the cortex should contribute more than the WM voxels inside the large WM regions since it is more effective to match WM voxels near the cortex than the inside regions.

To incorporate utility information into image registration procedures, we propose a novel image similarity measure, referred to as quantitative–qualitative measure of MI (Q-MI). Q-MI considers not only the probability of image intensity, but also the utility of each voxel, when calculating MI of two images. This is significantly different from the conventional MI measure that only considers the quantitative aspect of information based on the image intensity distribution. It is worth noting that the utility of each voxel in an image can be determined according to the regional saliency value calculated from the scale-space map of this image [27,28]. Therefore, the salient voxels will have higher utility values, and they will contribute more in measuring the MI of the two images under registration. That is, the voxels with high utilities play major roles in determining the transformation between the images.

Importantly, the utility values of voxels are not fixed, and they will be hierarchically updated during the registration procedure, with all voxels contributing equally in the final stage. In particular, the initial utility of each voxel will be assigned according to its saliency value [27,28]; with the progress of image registration, this utility will gradually move towards one. Thus, by mainly focusing on the voxels (or the regions) with higher utilities in the initial registration procedure, the robustness of registration can be improved. Also, by changing each joint utility to one in the final stage, the sub-voxel accuracy of registration can be retained as that obtained by the conventional MI-based registration methods, because of using MI in the final registration procedure. This hierarchical framework makes the Q-MI-based image registration not only robust but also accurate as demonstrated in the experiments.

The proposed Q-MI has been applied to the rigid registration of clinical brain images, such as MR, CT and PET images, obtained from the Retrospective Registration Evaluation Project (RREP) [29]. Experimental results show that, compared to conventional MI-based registration methods, the Q-MI-based registration method can provide a smoother registration function with a relatively larger capture range. It is also much more robust and has much higher success rates for the image registration.

The remainder of this paper is organized as follows. The definition of Q-MI is first provided in Section 2, and then the Q-MI-based registration method is described in Section 3. The performance of this Q-MI-based registration method is demonstrated in Section 4. This paper concludes in Section 5.

2. Quantitative–qualitative measure of MI (Q-MI)

In this section, the basic concepts of information and MI are first briefly introduced (Section 2.1). Then, the quantitative–qualitative measure of information (Section 2.2) and MI (Section 2.3) is presented.

2.1. Measure of information and MI

The information measure of an event is defined as a value related to the uncertainty or probability of occurrence of that event [30]. The self-information of an even E_n with probability p_n can be defined as

$$H(E_n) = -\log p_n. \tag{1}$$

For a set of events $E = \{E_1, \ldots, E_n, \ldots, E_N\}$ with respective probabilities $P = \{p_1, \ldots, p_n, \ldots, p_N\}$, the average information contained in E can be calculated by Shannon's entropy [31],

$$H(E) = H(E_1, \ldots, E_n, \ldots, E_N) = \sum_{n=1}^{N} p_n(-\log p_n). \tag{2}$$

In image processing, each intensity can be considered as an event, and a whole image can be considered as a set of such events; therefore, Shannon's entropy of an image reflects the complexity of intensity distribution in this image.

Since, in practice, we may not be able to obtain the true probabilities of occurrence of N events in E, i.e., $P = \{p_1, \ldots, p_n, \ldots, p_N\}$, but their estimates $Q = \{q_1, \ldots, q_n, \ldots, q_N\}$, the relative entropy, called Kullback–Leibler distance [30], can be used to measure how close Q approaches P,

$$D(E) = \sum_{n=1}^{N} p_n \left(\log \frac{p_n}{q_n}\right). \tag{3}$$

MI is used to measure the amount of information one set of events contains about another set of events. Given two sets of events, $E = \{E_1, \ldots, E_n, \ldots, E_N\}$ with a probability distribution $P = \{p_1, \ldots, p_n, \ldots, p_N\}$, and $F = \{F_1, \ldots, F_m, \ldots, F_M\}$ with a probability distribution $Q = \{q_1, \ldots, q_m, \ldots, q_M\}$, their MI is defined as the relative entropy between the joint distribution and the product of marginal distributions as follows:

$$MI(E, F) = \sum_{n=1}^{N}\sum_{m=1}^{M} p(E_n, F_m) \log \frac{p(E_n, F_m)}{p_n q_m}. \tag{4}$$

When events in E is independent of events in F, $p(E_n, F_m) = p_n q_m$, and $MI(E, F) = 0$.

MI can be applied to multimodality image registration, since it measures the amount of information that the two different

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

modality images of the same subject contain each other. Thus, by maximizing MI, a transformation can be obtained to align one image onto another.

2.2. Quantitative–qualitative measure of information

The above traditional information measures only define the quantitative aspect of information based on the probabilities of events. In fact, the occurrence of events often causes different influences and effects. Thus, besides the quantitative measures, the qualitative aspect of information should also be considered in order to fully justify the information of events, especially when we need to consider the effects of their occurrence.

In [32], Belis and Guiasu presented a quantitative–qualitative measure of information in cybernetic systems, which consists of two aspects of an event, i.e., the quantitative one related to its probability of occurrence, and the qualitative one related to its utility for the fulfillment of the goal. In particular, the quantitative–qualitative information supplied by event E_n with a probability p_n and a utility u_n is defined as

$$QH(E_n; u_n) = u_n(-\log p_n). \tag{5}$$

The utility u_n is a subjective notion directly connected to the goal to achieve. Generally, it can be set as any non-negative real number, and a larger value means the event is more useful. If the utility in Eq. (5) is set to one, Eq. (5) becomes Eq. (1).

Let $U = \{u_1, \ldots, u_n, \ldots, u_N\}$ be the utilities of the events in E, the quantitative–qualitative measure of information of E can be defined as follows:

$$QH(E; U) = \sum_{n=1}^{N} u_n p_n(-\log p_n). \tag{6}$$

When the utilities of all events are equal to one, Eq. (6) becomes Eq. (2). In practice, since the utilities of different events are different, $QH(E; U)$ does not achieve its maximum when all the events have uniform probabilities.

Based on the work of Belis and Guiasu [32], Taneja [33] presented a quantitative–qualitative measure of relative information as

$$QD(E; U) = \sum_{n=1}^{N} u_n p_n \log \frac{p_n}{q_n}. \tag{7}$$

Notice that $u_n \log(p_n/q_n) = u_n \log p_n - u_n \log q_n = -u_n \log q_n - (-u_n \log p_n)$, and $-u_n \log p_n$ is referred to as the *useful* self-information conveyed by an event with probability p_n and utility u_n; thus term $u_n \log(p_n/q_n)$ can be regarded as the *useful information gain*. When the utilities in Eq. (7) equal one, Eq. (7) becomes Eq. (3).

2.3. Quantitative–qualitative measure of MI

As mentioned in Section 2.2, Shannon's entropy-based MI is based on the probabilities of occurrence of events, and does not consider the particular qualitative aspects of events with respect to the goal, i.e., the utility of each event. In order to incorporate this qualitative aspect of event into the measure of MI, we define a quantitative–qualitative measure of MI (Q-MI) as follows:

$$QMI(E, F; U) = \sum_{n=1}^{N} \sum_{m=1}^{M} u(E_n, F_m) p(E_n, F_m) \log \frac{p(E_n, F_m)}{p_n q_m}, \tag{8}$$

where $u(E_n, F_m)$ represents the joint utility of events E_n and F_m. Like the definitions given in Eqs. (6) and (7), Q-MI focuses on the useful information that one set of events tells about another set of events. When all the joint utilities are set to one, Eq. (8) becomes the conventional MI. When the events in E are independent of the events in F, Q-MI is equal to zero, i.e., $QMI(E, F; U) = 0$. Moreover, when $u(E_n, F_m) = 0$ for all combinations of n and m, Q-MI is equal to zero. This means that the set of events E tells nothing useful about the set of events F.

It is worth noting that, although the utility of an event u_n in Eqs. (5)–(8) can be set between 0 and 1, the sum of utilities of all events, i.e., $\sum_{i=1}^{N} u_n$, is not necessarily to be 1. This is because the concept of utility is different from the concept of weight. The utility of an event just indicates the amount of contribution that this event provides for a given goal, whereas the weight is generally used to balance different terms. Moreover, increasing the utility of one event will not necessarily decrease the utilities of other events. Thus, Q-MI can not be simply considered as weighted MI.

3. Q-MI-based multimodality image registration

In this section, we will first describe the Q-MI-based registration algorithm by employing Q-MI as a similarity measure. Then, the method of computing the utility for each voxel in an image will be introduced (Section 3.2). Afterwards, the method of estimating the joint utility of two images will be presented, which will be used for the calculation of Q-MI (Section 3.3). Finally, the optimization method used in this registration algorithm will be briefly described (Section 3.4).

3.1. General description

As mentioned in Section 2.3, Q-MI gives the amount of useful information that one image contains in another image. Thus, maximizing the Q-MI of the two images under registration ensures the alignment of the two images mainly based on the voxels with useful information. The assignment of utility for each voxel in the images depends on the applications at hand, which will be discussed in Section 3.2.

The Q-MI-based registration algorithm can be formulated as follows. By denoting the two images under registration as the reference image R and the floating image F, and letting I_R and I_F be the intensity values of the two images, respectively, the goal of registration is to find out a transformation T so that the transformed floating image F_T can be well aligned with

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

the reference image R. The Q-MI between the reference image R and the transformed floating image F_T under the current transformation T can be derived from Eq. (8) as follows:

$$QMI(R, F_T; U) = \sum\sum u(I_R, I_{F_T}) p(I_R, I_{F_T}) \log \frac{p(I_R, I_{F_T})}{p(I_R) p(I_{F_T})}, \quad (9)$$

and the optimal transformation T^* between images R and F can be obtained by maximizing the Q-MI, i.e.,

$$T^* = \arg \max_T QMI(R, F_T; U). \quad (10)$$

In Eq. (9), the marginal and the joint image intensity distributions, i.e., $p(I_R)$, $p(I_{F_T})$ and $p(I_R, I_{F_T})$, can be estimated by simply calculating the marginal and the joint histograms of images within the overlap region. $u(I_R, I_{F_T})$ is a *joint utility* for each intensity pair (I_R, I_{F_T}), which will be determined by the utilities of both images in Section 3.3. Notice that, instead of fixing the joint utilities in the whole registration procedure, they are hierarchically updated at different registration stages, as described next.

In particular, each joint utility $u(I_R, I_{F_T})$ is changed from its initial value to one in the whole registration procedure. That means, initially the voxels with higher utility values contribute more on the alignment of two images; with the progress of registration, the utilities of voxels gradually change towards one, thus all voxels finally contribute equally for the image registration. Accordingly, this Q-MI-based registration algorithm not only improves the robustness of registration, but also retains the sub-voxel accuracy of registration as the conventional MI-based registration methods did, because of employing MI measure in the final registration stage. Specially, $u(I_R, I_{F_T})$ gradually changes with the registration stage t according to the function $u_0(I_R, I_{F_T}) + \alpha(t) \cdot (1 - u_0(I_R, I_{F_T}))$, where $\alpha(t)$ is a parameter increasing from 0 to 1 with the increase of t during the registration procedure, and $u_0(I_R, I_{F_T})$ is the initial joint utility which will be determined by Eq. (13) in the next. Thus, $\alpha(t)$ controls the changes of $u(I_R, I_{F_T})$ from its initial value $u_0(I_R, I_{F_T})$ to 1. As proved in our experiments, this Q-MI-based registration algorithm can converge, since it finally becomes MI-based registration method.

3.2. Utility of each voxel

Each voxel in the image is unique, and it has its own roles. One important difference between those roles is the amount of significance. For example, the voxels that lie in the region of interest or at the boundary of region of interest are more significant for image analysis and understanding task, compared to the voxels that lie in the background. Although many methods have been proposed to characterize important features of image voxels, how to characterize the properties of each voxel still remains a hot topic in computer vision and pattern recognition fields.

Many techniques have been developed to define the saliency of a voxel, i.e., using edges [34,35], corners [36] and key points [37–39]. Gradient operator is one simple image detector, and it is able to identify the location of intensity changes. As we indicated in the Introduction, gradient map has been widely incorporated into the MI-based registration methods. However, gradient is a local feature, and it is sensitive to noise. On the other hand, saliency measure [27], defined from scale-space map for each voxel in the image, is robust to noise and it considers regional information. Accordingly, the saliency definition is adopted here for representing the saliency of each voxel, and also the utility of this voxel in image registration.

Saliency measure [27] can be defined for each voxel in an image, based on local image complexity such as entropy. Basically, a point can have low image complexity if only a small region around this point is evaluated; however, it will have high image complexity when a large region is evaluated. Therefore, it is important to find the best scale such that entropy is maximized. On the other hand, local image exhibits self-similarity over a large range of scales. Therefore, self-dissimilarity measure should also be considered into the definition of saliency as in Ref. [27].

The best scale for a point can be determined by analyzing entropy in its local regions of different size. For each voxel x, we first calculate the probability distribution of intensity i, $p_i(s,x)$, in a spherical region of radius s centered at x. Then, we calculate the local entropy $L(s,x)$ from $p_i(s,x)$, as defined below:

$$L(s, x) = -\sum_i p_i(s, x) \log p_i(s, x). \quad (11)$$

The best scale s_x for the region centered at voxel x is selected as the one that maximizes the local entropy $L(s, x)$. Fig. 1 shows an example of finding the best scale for a given location. In Fig. 1(a), a green point denotes the point under consideration, while the red circles denote the different sized regions. Fig. 1(b) shows the change of local entropy versus scale, which can be used to determine the best scale for the considered point.

Since larger scale and higher local image difference are also preferred, the saliency value of voxel x, denoted as $A(s_x, x)$, is defined by the maximal local entropy value, weighted by both the best scale s_x and a self-dissimilarity measure in the scale space,

$$A(s_x, x) = L(s_x, x) \cdot \left\{ s_x \cdot \sum_i \left\| \frac{\partial p_i(s, x)}{\partial s} \Big|_{s_x} \right\| \cdot \Delta i \right\}, \quad (12)$$

where Δi is a constant, and the self-dissimilarity measure is defined as the sum of absolute differences of the probability distribution $p_i(s,x)$.

By measuring saliency over the whole image, each voxel has a saliency value to represent its significance in the image and also utility in image registration. By using this saliency measure as the utility of each voxel in image registration, Q-MI can provide more robust registration results, compared to the conventional MI-based registration methods, which treats each voxel equally.

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

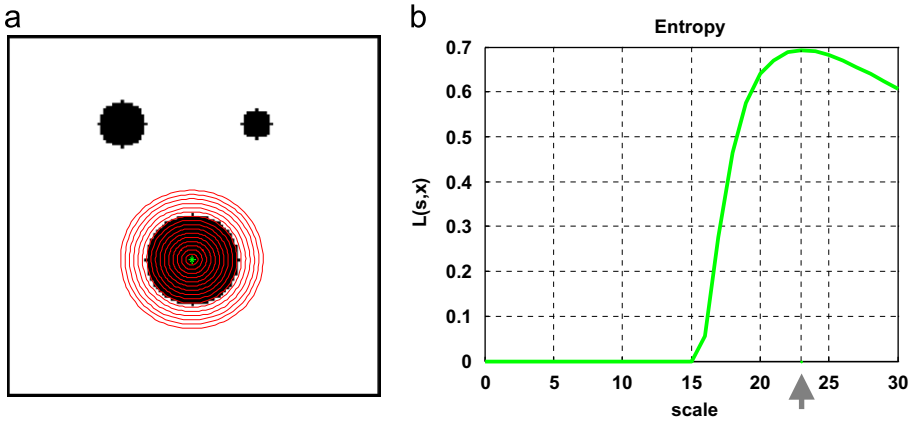


Fig. 1. An example of finding the best scale for a given point. (a) Red circles denote the different scales, (b) the curve shows the change of local entropy versus scale. The best scale for this example is about 23, as indicated by the gray arrow (notice the radius of the black region in (a) is 15).

3.3. Joint utility

Once the utility has been defined for each voxel in the two images under registration, we are ready to define the joint utility for each intensity pair in the two images. Let $I_R(x)$ be the intensity of reference image R at a location x, and $I_F(y)$ be the intensity of floating image F at a corresponding location y, let $A_R(x)$ and $A_F(y)$ be the utility values of voxel x in R and voxel y in F, respectively, the joint utility of intensity pair i and j can be defined as

$$u_0(i,j) = \sum_{\substack{x,y\in\Omega \\ I_R(x)=i, I_F(y)=j}} A_R(x) \bullet A_F(y), \tag{13}$$

where Ω is the overlap region of images R and F, and "$\cdot$" acts as an operator to combine the utility values from the reference image R and the floating image F. The voxel x in the reference image R and the voxel y in the floating image F are the corresponding points under the current transformation T. In order to emphasize the joint occurrence of two salient voxels, we use *multiplication* operation to integrate the saliency values from two images R and F.

Fig. 2 shows an example of estimating the joint utility of an intensity pair according to Eq. (13). Given a reference image (Fig. 2(a)) and a floating image (Fig. 2(b)), the joint utility of an intensity pair, e.g., (50, 222), can be estimated as follows. First, we find all the corresponding point pairs such as $\{(x_1, y_1), (x_2, y_2), \ldots, (x_5, y_5), \ldots\}$ in the two images, so that all the x points in Fig. 2(a) have intensity 50, and all the y points in the floating image have intensity 222. Then, the joint utility $u_0(50, 222)$ is calculated from these point pairs as demonstrated in Fig. 2(c).

As we have mentioned before, the joint utility is used to reflect the importance of intensity pairs under the current registration state, not used to balance the roles of different intensity pairs in the image registration. As such, the joint utility is not normalized when it is used to calculate Q-MI. Since we determine the joint utilities according to the utilities of voxels in the two images, Q-MI thus contains the amount of useful information that one image contains about another. Finding a transformation that maximizes the Q-MI of these two images will align the floating image onto the reference image, wherein the voxels with higher utilities act the major role in determining the transformation parameters.

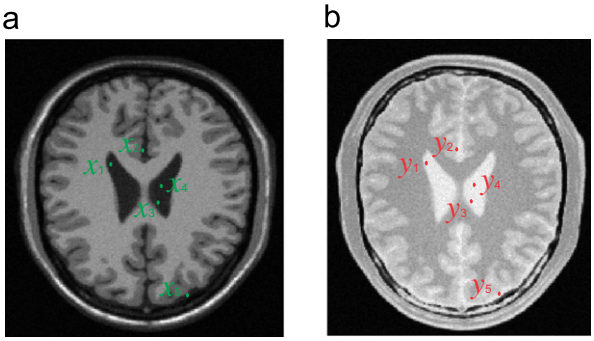
3.4. Optimization

To increase the robustness and also save the computation time, a multi-resolution framework of registration [40,41] is performed using four resolutions, as described next.

- In the coarsest resolution, an exhaust search over all the prescribed rotation parameters (including both coarse and fine rotation parameters) is performed. Coarse and fine rotation parameters are obtained by dividing the input search range by their own rotation increments [40,41]. For each coarse rotation parameter, an optimization of the translation parameters is performed; while for each fine rotation parameter, the translation parameters are first interpolated from the results using the coarse rotation parameters, and then optimized. The objective of these two steps is to try a set of rotation parameters and find out the best rotation and translation parameters. Since the calculation is carried out under the lowest resolution, for each best rotation parameter, a number of candidate transformation parameters in the local maxima are selected for further consideration in the next step.
- In the second coarsest resolution, a local optimization is performed on each of the candidate transformation parameters, with random perturbations. After optimization, a transformation parameter that yields the highest values for Q-MI is chosen, and it is optimized in the next step.
- In both the second finest and the finest resolutions, a local optimization is also performed on the input candidate transformation parameters, and then the final registration result is obtained.

It is worth noting that, in this paper, only rigid transformation with three rotation and three translation parameters are considered. Also, for the purposes of comparison, a similar

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002



c

	x_1	x_2	x_3	x_4	x_5
$I_R(x)$	50	50	50	50	50
$A_R(x)$	1.23	1.35	1.16	1.21	0.92

	y_1	y_2	y_3	y_4	y_5
$I_F(y)$	222	222	222	222	222
$A_F(y)$	1.12	1.16	0.97	0.96	1.18

$$u_0(50,222) = 1.23*1.12 + 1.35*1.16 + 1.16*0.97 + 1.21*0.96 + 0.92 + 1.18 = 6.41$$

Fig. 2. An example of estimating joint utility of an intensity pair. (a) A reference image, (b) a floating image, and five corresponding point pairs with intensity 50 in image (a) and intensity 222 in image (b) are selected, (c) a procedure of demonstrating the estimation of joint utility of the intensity pair (50, 222), from five selected point pairs.

optimization scheme is performed for the MI-based registration algorithms.

4. Experimental results

A number of experiments have been carried out to evaluate the performance of the Q-MI-based registration algorithm in aligning the multimodal images of brains such as MR PD-weighted, MR T1-weighted, MR T2-weighted, CT, and PET images, obtained from the Vanderbilt Retrospective Registration Project [29]. MR images are with 256×256 voxels in plane, 20–16 slices, and voxel size of 1.25 mm $\times$ 1.25 mm $\times$ 4.0 mm. CT images are with 512×512 voxels in plane, 27–34 slices, and voxel size of 0.65 mm $\times$ 0.65 mm $\times$ 4.0 mm. PET images has 15 slices, each with 128×128 voxels and the voxel size of 2.59 mm $\times$ 2.59 mm $\times$ 8.0 mm. It is worth noting that there exist scale differences between these modality images due to different voxel sizes used. In the following, we first compare the Q-MI- and MI-based registration functions, and then evaluate the robustness and accuracy of the Q-MI- and MI-based registration algorithms, respectively.

4.1. Registration functions

An ideal registration function that measures the similarity of two images should be smooth and convex *with respect to* different transformation parameters. Also, the global maximum of the registration function should be close to the correct transformation parameters that align two images perfectly. Moreover, the capture range around the global maximum should be as large as possible, and the number of local maxima of the registration function should be as small as possible. These criteria will be used to evaluate the registration functions generated by the Q-MI and the MI measurements, respectively.

The registration function of Q-MI can be generated by measuring the Q-MI of two images under all possible transformations. For the images used in our experiments, their relative transformation parameters can be determined with the aid of four fiducial markers implanted in the patients [29]. That means, we can first align all testing images into the same space by using the four fiducial markers, and then calculate the Q-MI measures of two testing images under different transformations (i.e., rotation or translation), thereby obtaining a registration function of Q-MI. Similarly, a registration function of MI can be obtained.

The registration functions of Q-MI and MI are extensively compared on three different cases, i.e., registering MR T1 image with (1) MR T2 image, (2) CT image, and (3) PET image. For visual inspection, the brain images used in these experiments are shown in Figs. 3(a)–(d), along with their color-coded utilities in Figs. 3(e)–(h). Notice that all registration functions in these experiments are obtained by using the down-sampled images with scale 3 at each dimension, and also the original joint utility $u_0(I_R, I_{F_T})$ is used to calculate the registration function of Q-MI. Experimental results on these three different registration cases are detailed below. Notice that each registration function in Figs. 4–7 represents a single Q-MI function of two images under registration, with respect to each possible translation or rotation existing between them.

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

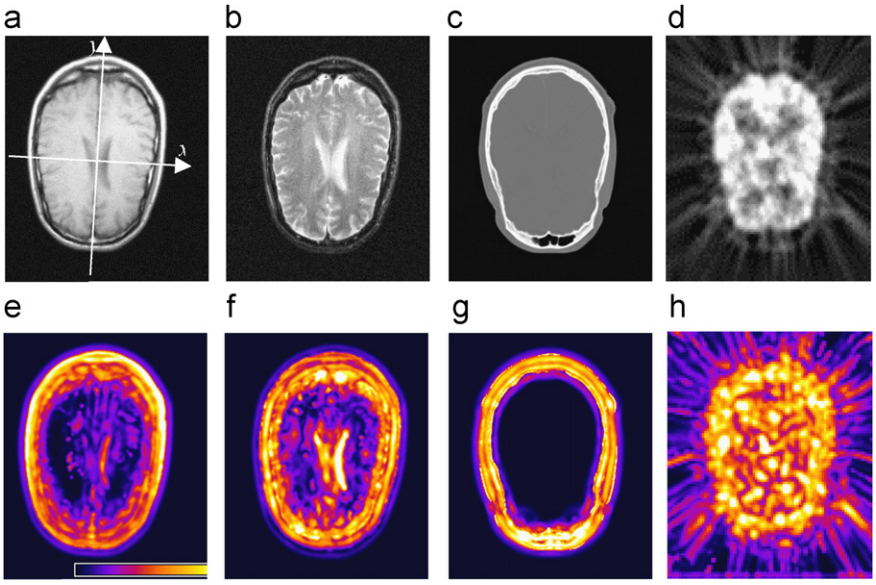


Fig. 3. Different modality images of the same brain, used for evaluating the registration functions of Q-MI and MI. (a) MR T1 image, (b) MR T2 image, (c) CT image, (d) PET image, and (e)–(h) the respective color-coded utility maps of the images in (a)–(d). Here, yellow and white denote the high utilities.

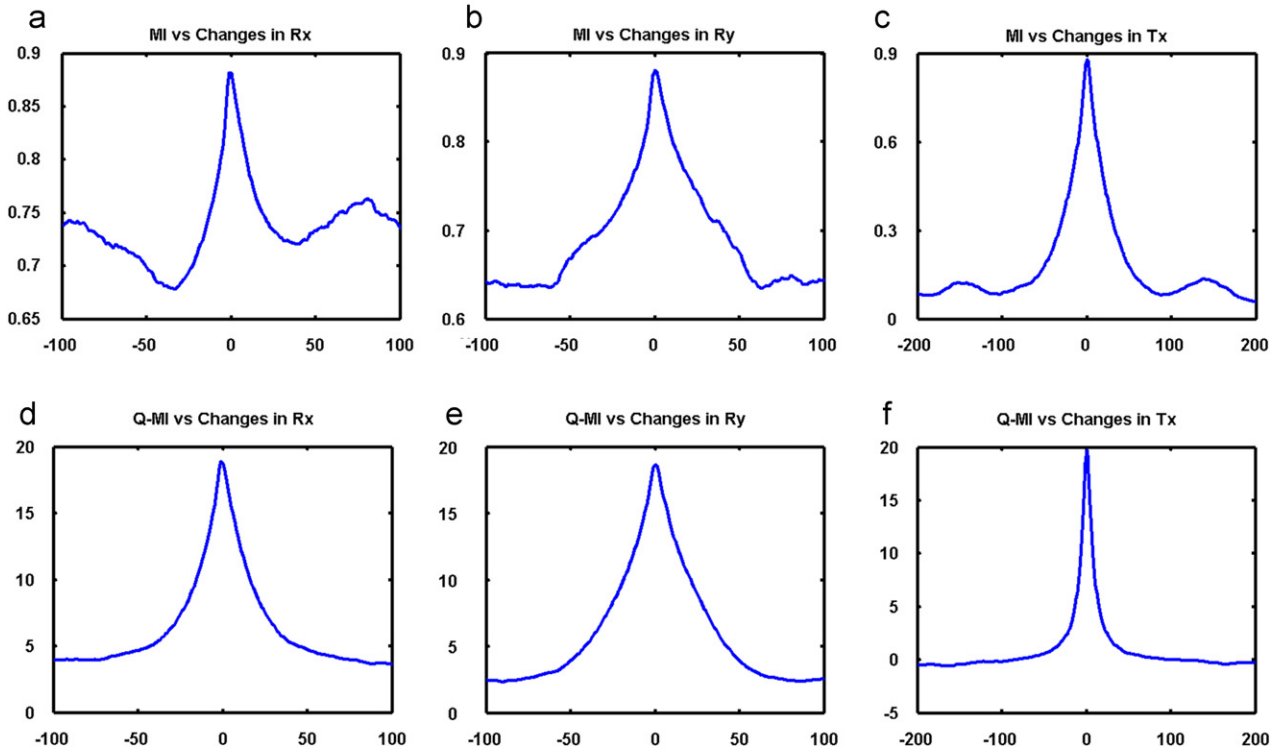


Fig. 4. The registration functions of Q-MI and MI in aligning MR T1 image with MR T2 image. The top row shows the registration functions of MI, while the bottom row shows the registration functions of Q-MI. From left to right: (a) and (d) rotation around x-axis, (b) and (e) rotation around y-axis, (c) and (f) translation along x-axis.

4.1.1. Registering MR T1 image with MR T2 image

MR T1 and T2 images are both structural images, containing detailed anatomical information. The registration between them is relatively easy. The respective registration functions of Q-MI and MI *with respect to* different rotations and translations are plotted in Fig. 4. The top row shows the registration functions of MI versus the rotation around x-axis (R_x), the rotation around y-axis (R_y), and the translation along x-axis (T_x), respectively. The bottom row shows the corresponding registration functions of Q-MI.

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

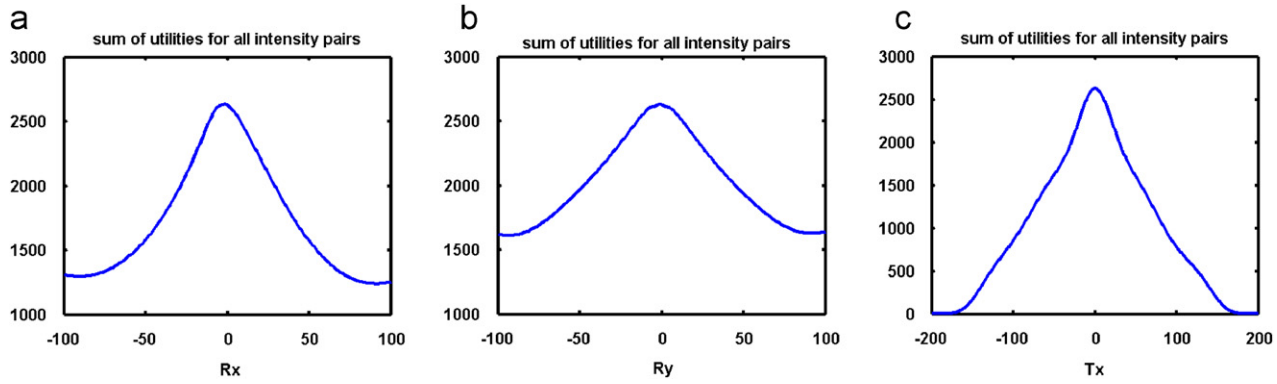


Fig. 5. The total amount of joint utilities within the overlap region of the two images, for the cases of (a) rotation around x-axis, (b) rotation around y-axis, and (c) translation along x-axis, respectively.

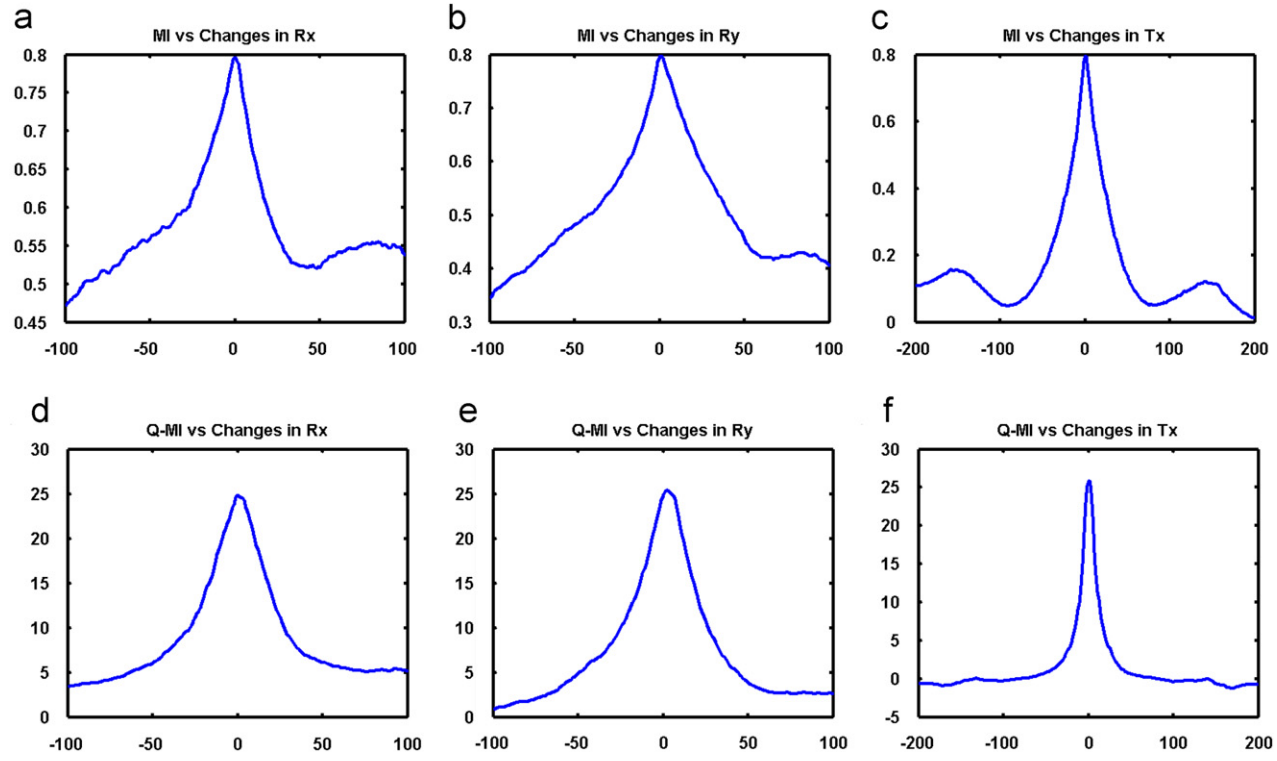


Fig. 6. The registration functions of Q-MI and MI in aligning MR T1 image with CT image. The top row shows the registration functions of MI, while the bottom row shows the registration functions of Q-MI. From left to right: (a) and (d) rotation around x-axis, (b) and (e) rotation around y-axis, (c) and (f) translation along x-axis.

By comparing (a) and (b) with (d) and (e) in Fig. 4, respectively, we can observe that the registration functions of Q-MI is much smoother than those of MI. Moreover, the capture range in the registration function of Q-MI is relatively large, e.g., the change of registration function with respect to the rotations around x- and y-axes is smoothly extended relatively far from the global maximum. All of these excellent properties should contribute to the integration of utility information into the Q-MI calculation, particularly with salient voxels contributing more in the image registration. Notice that the absolute value of Q-MI is larger than that of MI. This is because the joint utility used for each intensity pair is not normalized by the total utility in the whole image.

By comparing Figs. 4(c) with (f), we can observe that the registration function of Q-MI is smoother than that of MI, but it drops quickly with the increase in translation, thus it does not extend its capture range like it does in rotation. This is because the overlay region of the two images drops very quickly when translating one image from the other, compared to the cases of rotating images. This makes the total amount of joint utilities in the overlap region of the two images also drops very quickly, as proved in Fig. 5. According to Fig. 5(c), the total amount of

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

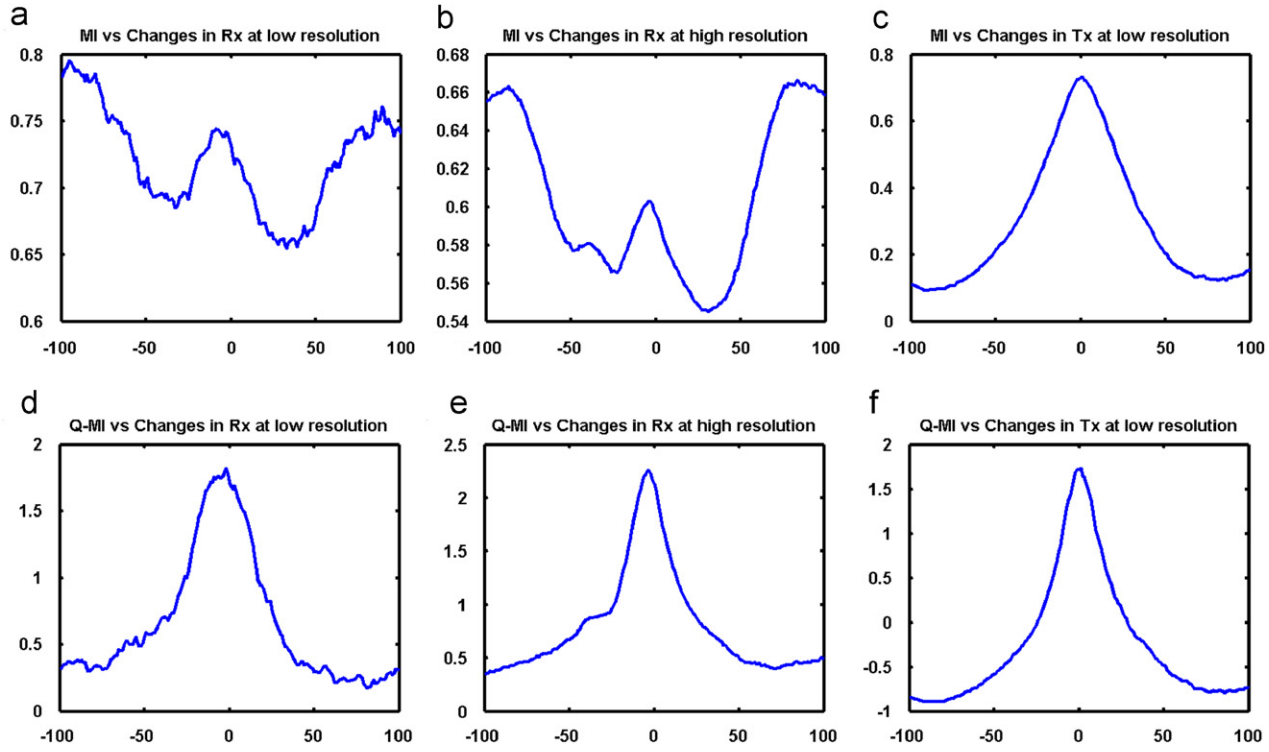


Fig. 7. The registration functions of Q-MI and MI in aligning MR T1 image with PET image. The top row shows the registration functions of MI, while the bottom row shows the registration functions of Q-MI. From left to right: (a) and (d) rotation around x-axis at low resolution, (b) and (e) rotation around x-axis at high resolution, (c) and (f) translation along y-axis at low resolution.

joint utilities becomes zero, when the translation is sufficiently large and thus no overlap region exists. On the other hand, even the rotation can be very large, the overlap region between two brain images always exist, thus the total amount of joint utilities will not decrease to zero, as indicated in Figs. 5(a) and (b).

4.1.2. Registering MR T1 image with CT image

Although MR T1 and CT images are quite different, they still have similar structures, e.g., skull. According to the registration functions of Q-MI and MI shown in Fig. 6, it can be observed again that the registration functions of Q-MI are much smoother and have relatively large capture range, compared to the registration functions of MI, for the cases of rotating images around x- and y-axes. Notice that, for the case of translation, although Q-MI does not extend its capture range, it still makes the registration function much smoother.

4.1.3. Registering MR T1 image with PET image

MR T1 image is different from PET image: one is a structural image, while the other is a functional image usually in very low resolution. The only common information in these two modality images is the overall shape of the brain, therefore registering these two images becomes relatively difficult. According to the registration functions of Q-MI and MI shown in Fig. 7, we can see again that the registration functions of Q-MI are relatively smooth, although not as smooth as the registration functions generated when aligning two *structural* images as demonstrated before.

4.2. Robustness of registration

A robust registration algorithm should be able to recover the true transformation between the two images under registration, even if the initial misalignment between them is large. Accordingly, the robustness of the Q-MI- and MI-based registration algorithms is evaluated on various amounts of initial misalignment between MR T1 image and MR T2 image, as well as between MR T1 image and CT image.

Four sets of tests have been performed, with the degrees of initial misalignment rotation angles randomly picked from the four different rotation ranges, i.e., $[-5, 5]$, $[-10, 10]$, $[-20, 20]$, and $[-30, 30]$ degrees, respectively. In particular, altogether, we performed 50 tests for each of four different rotation ranges. Notice that, for each test, the registration is considered as successful if the difference between the estimated transformation and the ground-truth transformation is less than a pre-defined threshold. Similar to Ref. [42], the threshold for each transformation parameter can be, respectively, selected as 4° for rotation around x-axis or y-axis, 2° for rotation around z-axis, 2 mm for translation in x-axis or y-axis, and 3 mm for translation in z-axis.

The respective numbers of successful registration cases for each of the four different rotation ranges, using the Q-MI- and MI-based registration algorithms are summarized in Fig. 8. These results indicate that the Q-MI-based registration algorithm produces much higher successful rate, and thus it is more robust than the MI-based registration algorithm,

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

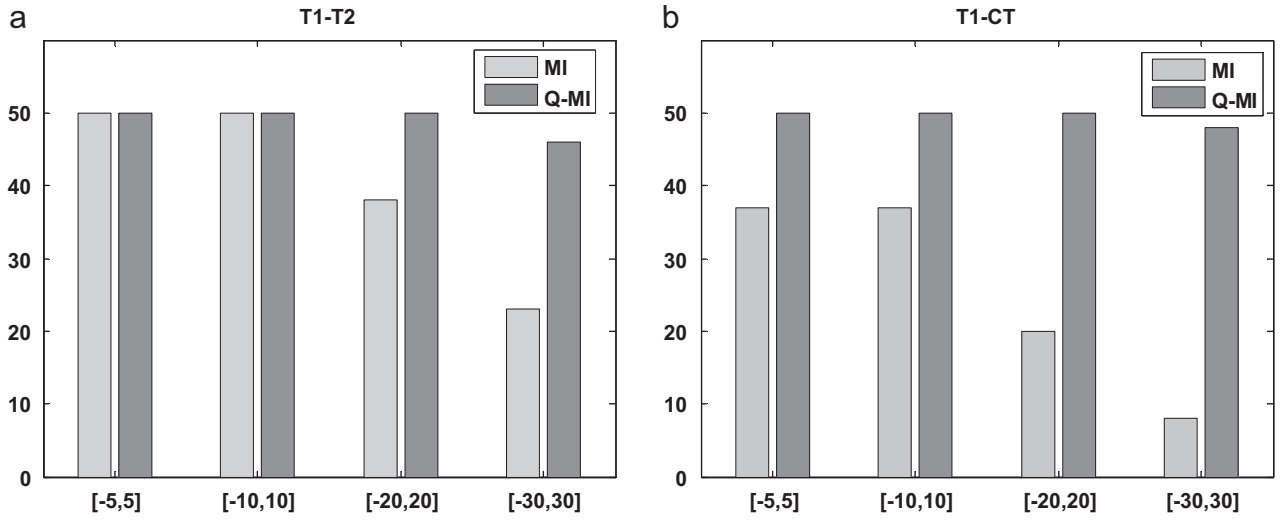


Fig. 8. The number of successful registrations for 50 tests of each of four different rotation ranges, using the Q-MI- and MI-based algorithms, respectively. (a) The results in registering MR T1 image with MR T2 image, (b) the results in registering MR T1 image with CT image.

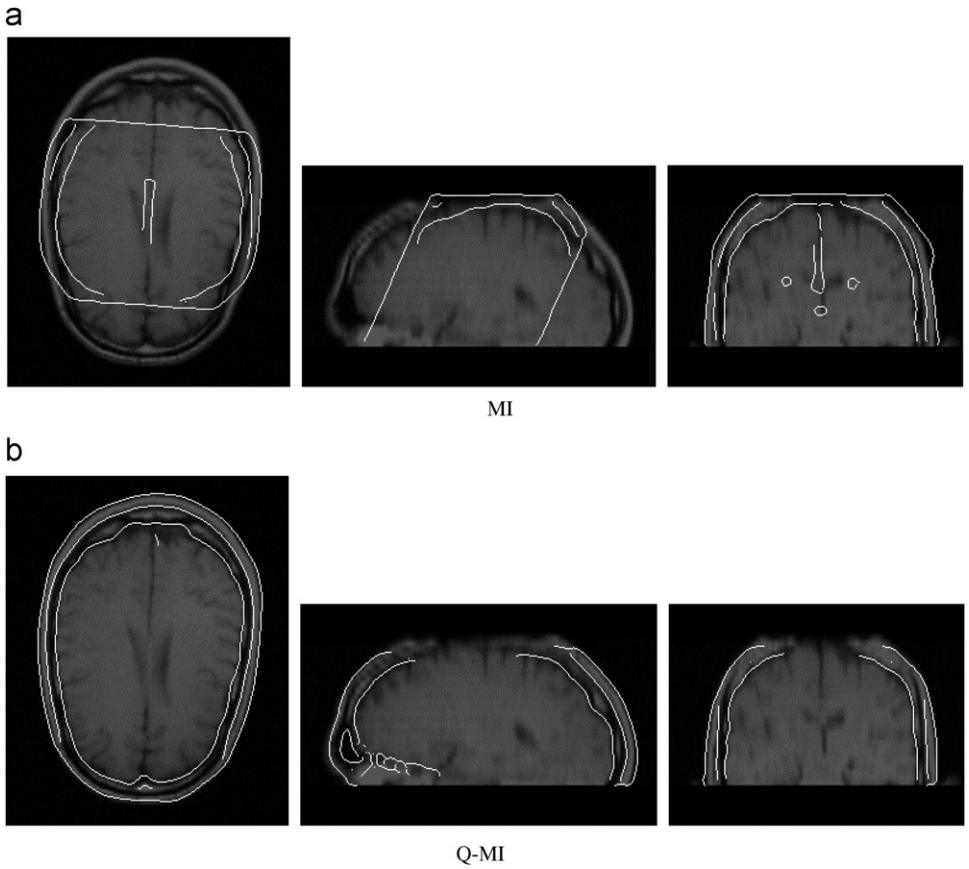


Fig. 9. Visual inspection for the registration results obtained using Q-MI- and MI-based registration algorithms, respectively. (a) A failing case by the MI-based registration algorithm, (b) the result for the same case by the Q-MI-based registration algorithm. The white contours are the boundaries of the aligned CT image, overlaid on the MR T1 image at three different cross-sectional views.

particularly for the fourth rotation range $[-30, 30]$, where the degrees of initial misalignment rotation around x-, y- and z-axis are randomly picked up from a relatively large rotation range. This phenomenon becomes much more obvious when registering the MR T1 image with the CT image as shown in Fig. 8(b). For example, the MI-based registration algorithm can be successful only for 8 out of 50 tests for the fourth rotation range $[-30, 30]$; on the contrary, the number of successful

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

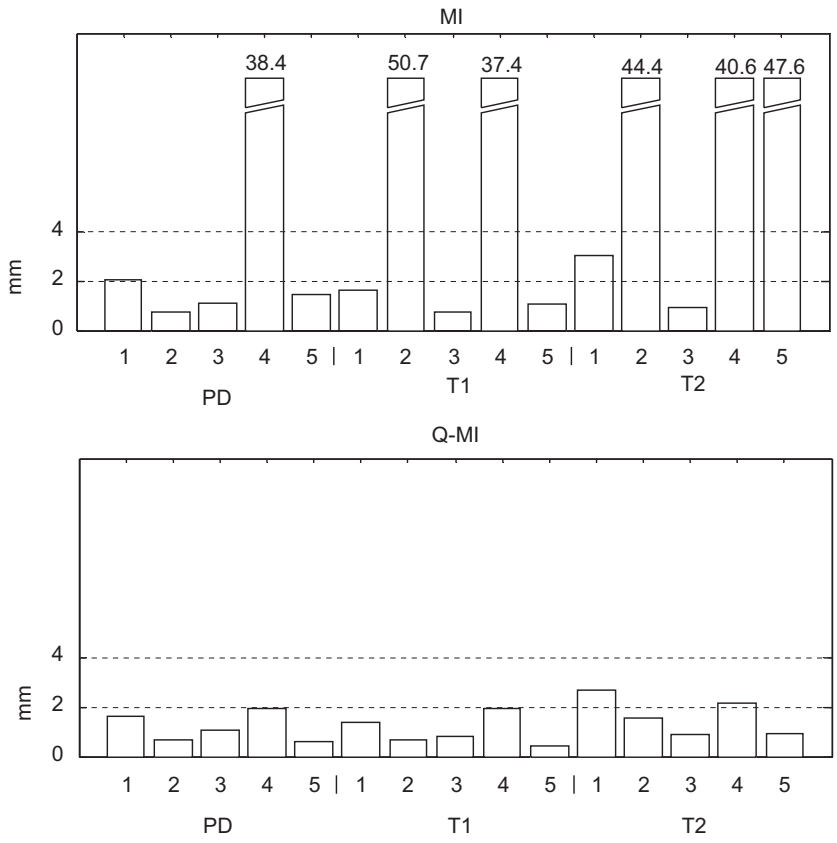


Fig. 10. Median alignment errors by the MI-based registration algorithm (top) and the Q-MI-based registration algorithm (bottom), when aligning CT images with MR (PD, T1 and T2) images.

registration of the Q-MI-based registration algorithm is close to 50. For visual inspection, one failing case by the MI-based registration algorithm is displayed in Fig. 9(a), while the result on the same case by the Q-MI-based registration algorithm is given in Fig. 9(b). Notice that the white contours in this figure represent the boundaries of the aligned CT image.

4.3. Registration accuracy

The registration accuracy between the Q-MI- and MI-based registration algorithms is also compared by aligning real CT images (or PET images) with MR images of the same subject, as demonstrated in Figs. 10 and 11 with provided median alignment errors. As indicated in Fig. 10, there are five cases for each experiment of registering CT images with MR (PD, T1 or T2) images. It can be observed that the Q-MI-based registration algorithm can successfully register all pairs of images, if 4 mm registration errors are regarded as successful. Also, for all of the successful registration cases, the accuracy of registration produced by the Q-MI-based registration algorithm seems a little bit better than that obtained by the MI-based registration algorithm. We will quantitatively compare the registration accuracy next.

As shown in Fig. 11, there are three cases for each experiment of registering PET images with MR (PD, T1 or T2) images. It can be also observed that the Q-MI-based registration algorithm is more successful, compared to the MI-based registration algorithm. However, there is still a case that is obviously failed. This might be due to the low resolution of PET images, and also strong noises in PET images. On the other hand, the alignment results by the MI-based registration algorithm become much worse when registering PET images with MR images, i.e., two out of the three tests have been failed. This is because of a lot of local maxima in the registration function generated by the MI-based registration algorithm, as indicated in Figs. 7(a) and (b). Notice that, for the successful registration cases, the accuracy of MI-based registration algorithm seems better than that by the Q-MI-based registration algorithm. The quantitative comparison is provided next.

For quantitatively comparing the registration accuracy of both algorithms, we calculate the average of the registration errors for all the successful cases when registering CT images with MR images, and PET images with MR images. As summarized in Table 1, the Q-MI-based registration algorithm can register CT images with MR images a little better than the MI-based registration algorithm. However, it becomes worse when registering PET images with MR images. Also, for both algorithms, the registration errors are smaller when registering CT images with MR images, compared to the cases of registering PET images with MR images. This is because CT and MR images are both structural images, thus they are relatively easy to register. It should be emphasized that, here, only the successful registration cases are compared. As indicated before, the

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

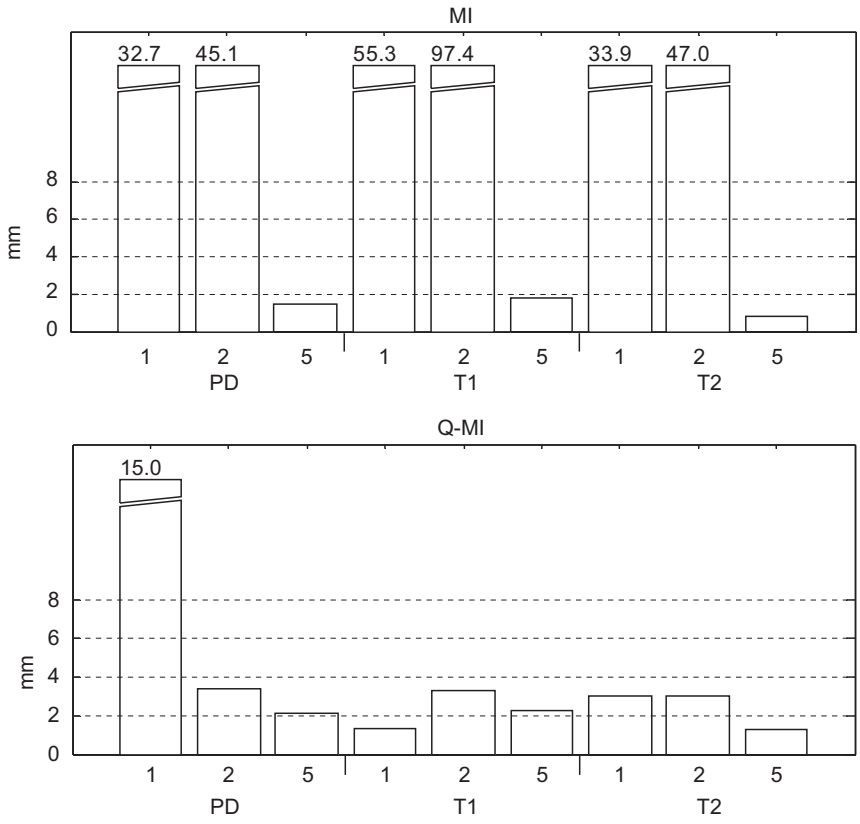


Fig. 11. Median alignment errors by the MI-based registration algorithm (top) and the Q-MI-based registration algorithm (bottom), when aligning PET images with MR (PD, T1 and T2) images.

Table 1
Comparison on the registration errors between the Q-MI and MI-based registration algorithms, on all successful registration cases

Image pair	MI-based registration (mm)	Q-MI-based registration (mm)
CT and MR	1.10	1.07
PET and MR	1.48	2.64

Q-MI-based registration algorithm is much more successful in image registration.

From Table 1, it can also be observed that the Q-MI-based registration method can achieve sub-voxel accuracy of registration. For CT and MR image registration, all registration errors are within 3 mm, which is reasonable considering the slice thickness of the CT images 4 mm. For PET and MR image registration, all registration errors are within 4 mm for successful cases, which is also reasonable considering the slice thickness of the PET images 8 mm.

5. Conclusion

A novel image similarity measure, called quantitative–qualitative measure of mutual information (Q-MI), has been presented for robust registration of multimodality brain images. By utilizing the concept of both quantitative and qualitative information measures of events, Q-MI incorporates utility information into the similarity measure of the two images, and hence it allows the registration procedure focusing more on matching the voxels with higher utility values, such as the regions of interest or salient voxels. Experimental results demonstrate that the registration function generated by Q-MI is much smoother than that by MI, and it has a larger capture range due to the incorporation of the joint utilities of the two images into the Q-MI measurement. Moreover, Q-MI-based registration algorithm is more robust than the MI-based registration algorithm in registering multimodality brain images. Finally, experimental results also indicate that the increase of robustness by our Q-MI-based registration will not sacrifice the accuracy of image registration, i.e., sub-voxel accuracy can be still achieved. In the future, we will extend this Q-MI-based registration algorithm to non-rigid image registration [43].

References

[1] Z. Lao, D. Shen, A. Jawad, B. Karacali, D. Liu, E. Melhem, N. Bryan, C. Davatzikos, Automated segmentation of white matter lesions in 3D brain MR images using multivariate pattern classification, in: Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging (ISBI), 2006.

[2] N. Hata, T. Dohi, S.K. Warfield, W.M. Wells, R. Kikinis, F.A. Jolesz, Multimodality deformable registration of pre- and intraoperative images for MRI-guided brain surgery, in: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 1998, pp. 1067–1074.

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

[3] A. Bharatha, M. Hirose, N. Hata, S.K. Warfield, M. Ferrant, K.H. Zou, E. Suarez-Santana, J. Ruiz-Alzola, A. D'Amico, R.A. Cormack, R. Kikinis, F.A. Jolesz, C.M.C. Tempany, Evaluation of three-dimensional finite element-based deformable registration of pre- and intraoperative prostate imaging, Med. Phys. 28 (12) (2001) 2551–2560.
[4] B. Zitova, J. Flusser, Image registration methods: a survey, Image Vision Comput. 21 (2003) 977–1000.
[5] A.W. Toga, P. Thompson, The role of image registration in brain mapping, Image Vision Comput. 19 (1–2) (2001) 3–24.
[6] C. Studholme, D.L.G. Hill, D.J. Hawkes, Multiresolution voxel similarity measures for MR-PET registration, in: Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI), vol. 3, 1995, pp. 287–298.
[7] F. Maes, D. Vandermeulen, P. Suetens, Medical image registration using mutual information, Proc. IEEE 91 (2003) 1699–1722.
[8] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, G. Marchal, Automated multi-modality image registration based on information theory, in: Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI), 1995, pp. 263–274.
[9] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, R. Kiknis, Multi-modal volume registration by maximization of mutual information, Med. Image Anal. 1 (1996) 35–51.
[10] J.P.W. Pluim, J.B. Maintz, M.A. Viergever, Mutual-information-based registration of medical images: a survey, IEEE Trans. Med. Imaging 22 (2003) 986–1004.
[11] D.J. Pettey, J.C. Gee, Using a linear diagnostic function and non-rigid registration to search for morphological differences between populations: an example involving the male and female corpus callosum, in: Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI), 2001, pp. 372–379.
[12] C.E. Rodriguez-Carranza, M.H. Loew, A weighted and deterministic entropy measure for image registration using mutual information, in: Proceedings of SPIE, Medical Imaging: Image Processing, vol. 3338, 1998, pp. 155–166.
[13] C.E. Rodriguez-Carranza, M.H. Loew, Global optimization of weighted mutual information for multi-modality image registration, in: Proceedings of SPIE, Medical Imaging: Image Processing, vol. 3661, 1999, pp. 89–96.
[14] B. Pompe, P. Blidh, D. Hoyer, M. Eiselt, Using mutual information to measure coupling in the cardiorespiratory system, IEEE Eng. Med. Biol. Mag. 17 (6) (1998) 32–39.
[15] C. Studholme, D.L.G. Hill, D.J. Hawkes, A normalized entropy measure for multi-modality image alignment, in: Proceedings of SPIE, Medical Imaging: Image Processing, vol. 3338, 1998, pp. 132–143.
[16] C. Studholme, D.L.G. Hill, D.J. Hawkes, An overlap invarivant entropy measure of 3-D medical image alignment, Pattern Recognition 32 (1) (1999) 71–86.
[17] D.L.G. Hill, C.R. Maurer Jr., R.J. Maciunas, J.A. Barwise, J.M. Fitzpatrick, M.Y. Wang, Measurement of intraoperative brain surface deformation under a craniotomy, Neurosurgery 43 (1998) 514–526.
[18] B. Likar, F. Pernus, A hierarchical approach to elastic registration based on mutual information, Image Vision Comput. 19 (2001) 33–44.
[19] C. Studholme, R.T. Constable, J.S. Duncan, Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model, IEEE Trans. Med. Imaging 19 (2000) 1115–1127.
[20] J.P.W. Pluim, J.B. Maintz, M.A. Viergever, Image registration by maximization of combined mutual information and gradient information, IEEE Trans. Med. Imaging 19 (2000) 809–814.
[21] D. Rueckert, M.J. Clarkson, D.L.G. Hill, D.J. Hawkes, Non-rigid registration using higher-order mutual information, in: Proceedings of SPIE, Medical Imaging: Image Processing, vol. 3979, 2000, pp. 438–447.
[22] C. Studholme, D.L.G. Hill, D.J. Hawkes, Incorporating connected region labeling into automated image registration using mutual information, in: Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 1996, pp. 23–31.
[23] D.B. Russakoff, C. Tomasi, T. Rohlfing, C.R. Maurer Jr., Image similarity using mutual information of regions, in: Proceedings European Conference on Computer Vision (ECCV), vol. 3023, 2004, pp. 596–607.
[24] M. Holden, L.D. Griffin, N. Saeed, D.L.G. Hill, Multi-channel mutual information using scale space, in: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2004, pp. 797–804.
[25] R. Gan, A. Chung, Multi-dimensional mutual information based robust image registration using maximum distance-gradient-magnitude, in: Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI), 2005.
[26] D. Shen, C. Davatzikos, HAMMER: hierarchical attribute matching mechanism for elastic registration, IEEE Trans. Med. Imaging 21 (11) (2002) 1421–1439.
[27] T. Kadir, M. Brady, Saliency, scale and image descriptions, Int. J. Comput. Vision 45 (2001) 83–105.
[28] X. Huang, Y. Sun, D. Metaxas, F. Sauer, C. Xu, Hybrid image registration based on configural matching of scale-invariant salient region features, in: Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop, 2004.
[29] J. West, J.M. Fitzpatrick, M.Y. Wang, B.M. Dawant, C.R. Maurer Jr., R.M. Kessler, R.J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P.A. van den Elsen, S. Napel, T.S. Sumanaweera, B. Harkness, P.F. Hemler, D.L.G. Hill, D.J. Hawkes, C. Studholme, J.B. Maintz, M.A. Viergever, G. Malandain, R.P. Woods, Comparison and evaluation of retrospective intermodality brain image registration techniques, J. Comput. Assist. Tomogr. (4) (1997) 554–566.
[30] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[31] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423.
[32] M. Belis, S. Guiasu, A quantitative–qualitative measure of information in cybernetic systems, IEEE Trans. Inform. Theory 14 (1968) 593–594.
[33] H.C. Taneja, R.K. Tuteja, Characterization of a quantitative–qualitative measure of relative information, Inform. Sci. 33 (1984) 217–222.
[34] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 8 (6) (1986) 679–698.
[35] F. Bergholm, Edge focusing, IEEE Trans. Pattern Anal. Mach. Intell. 9 (6) (1987) 726–741.
[36] R. Deriche, G. Giraudon, A computational approach for corner and vertex detection, Int. J. Comput. Vision 10 (2) (1993) 101–124.
[37] D.G. Lowe, Object recognition from local scale-invariant features, in: Proceedings of the International Conference on Computer Vision (ICCV), 1999, pp. 1150–1157.
[38] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vision 60 (2) (2004) 91–110.
[39] Y. Ke, R. Sukthankar, PCA-SIFT: a more distinctive representation for local image descriptors, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2004, pp. 506–513.
[40] M. Jenkinson, S. Smith, A global optimization method for robust affine registration of brain images, Med. Image Anal. 5 (2001) 143–156.
[41] M. Jenkinson, P. Bannister, M. Brady, S. Smith, Improved optimization for the robust and accurate linear registration and motion correction of brain images, NeuroImage 17 (2002) 825–841.
[42] Y.-M. Zhu, S.M. Cochoff, Likelihood maximization approach to image registration, IEEE Trans. Image Process. 11 (12) (2002) 1417–1426.
[43] C. Studholme, C. Drapaca, B. Iordanova, V. Cardenas, Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change, IEEE Trans. Med. Imaging 25 (5) (2006) 626–639.

About the Author—HONGXIA LUAN received her Bachelor degree from Xidian University and Master degree from Hangdian University. Presently, she is a Ph.D. candidate at Shanghai Jiao Tong University. Her research interests include medical image analysis and pattern recognition.

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002

About the Author—FEIHU QI was born in 1938. He is a Professor at Shanghai Jiao Tong University. He has published over 200 articles in various journals and proceedings of international conferences. His research interests include image processing, medical image analysis, pattern recognition, compute vision and artificial neural networks.

About the Author—ZHONG XUE received B.E. and M.E. degrees from Xidian University and Ph.D. degree from Singapore Nanyang Technological University. He is currently a Research Associate at University of Pennsylvania. He has published 33 international journal and conference papers in computer vision, image processing, and medical image analysis.

About the Author—LIYA CHEN received all her degrees (B.S., M.S., Ph.D.) from Shanghai Jiao Tong University. She has been an Associate Professor at the Department of Electronic Engineering of Shanghai Jiao Tong University since 2003. She has published over 20 articles in journals and proceedings of international conferences. Her research interests include image processing, medical image analysis, computer vision and artificial neural networks.

About the Author—DINGGANG SHEN received all of his degrees from Shanghai JiaoTong University. He is an assistant professor (tenure-track) in the Department of Radiology at University of Pennsylvania (Upenn) since July 2002. Before moving to Upenn, he was a faculty member in Johns Hopkins University. Dr. Shen is on the Editorial Board of Pattern Recognition, and served as a reviewer for numerous international journals and conferences. He has published over 160 articles in journals and proceedings of international conferences. His research interests include medical image analysis, pattern recognition, and computer vision.

Please cite this article as: H. Luan, et al., Multimodality image registration by maximization of quantitative–qualitative measure of mutual information, Pattern Recognition (2007), doi: 10.1016/j.patcog.2007.04.002