Maximum distance-gradient for robust image registration

Rui Gan, Albert C.S. Chung *, Shu Liao

Lo Kwee-Seong Medical Image Analysis Laboratory, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

Abstract

To make up for the lack of concern on the spatial information in the conventional mutual information based image registration framework, this paper designs a novel spatial feature field, namely the maximum distance-gradient (MDG) vector field, for registration tasks. It encodes both the local edge information and globally defined spatial information related to the intensity difference, the distance, and the direction of a voxel to a MDG source point. A novel similarity measure is proposed as the combination of the multi-dimensional mutual information and an angle measure on the MDG vector field. This measure integrates both the magnitude and orientation information of the MDG vector field into the image registration process.

Experimental results on clinical 3D CT and T1-weighted MR image volumes show that, as compared with the conventional mutual information based method and two of its adaptations incorporating spatial information, the proposed method can give longer capture ranges at different image resolutions. This leads to more robust registrations. Around 2000 randomized rigid registration experiments demonstrate that our method consistently gives much higher success rates than the aforementioned three related methods. Moreover, it is shown that the registration accuracy of our method is high.
© 2008 Elsevier B.V. All rights reserved.

Keywords: Image registration; Multi-modality; Medical imaging; Maximum distance-gradient; Mutual information

1. Introduction

A crucial element in the image registration process is a similarity measure to determine how well the images match with each other through a hypothesized spatial transformation. Mutual information, a useful concept from information theory, has been one of the most extensively adopted measures. It was originally introduced as a similarity measure on image intensity. Generally promising results have shown that mutual information is well-suited for multi-modal image registration (Maes et al., 1997; Meyer et al., 1997; Studholme et al., 1997; Wells et al., 1996; West et al., 1997). An important and beneficial property of mutual information is that it measures the statistical dependence (or in other words, shared information) between two image intensities, and does not assume any limiting (for instance, linear) relationship between image intensities in the images.

However, it has been suggested that the conventional mutual information based registration can result in misalignment in some cases (Penney et al., 1998; Pluim et al., 2000) and thus room for improvement exists. The conventional mutual information measure takes only image intensities into account, since by definition, it depends on the relative occurrence of intensities in individual images and co-occurrence of intensity pairs in both images. A known deficiency of mutual information is the lack of concern for any (local or global) spatial information, which may be present in individual images to be registered (Pluim et al., 2003; Rueckert et al., 2000). In order to increase registration performance, several researchers have proposed adaptations of the mutual information based registration framework to incorporate spatial information from individual images, for instance:

---

* Corresponding author. Tel.: +852 2358 7000; fax: +852 2358 1477.
 E-mail addresses: raygan@cse.ust.hk (R. Gan), achung@cse.ust.hk (A.C.S. Chung), liaoshu@cse.ust.hk (S. Liao).

Spatial feature vectors: Butz and Thiran (2001) applied mutual information to edge measure (e.g. gradient magnitude) space, which aligned object surfaces in the images to be registered. However, mutual information measure on edge features is sensitive to the sparseness of joint edge feature histograms. This may increase the difficulty of the optimization procedure.

By multiplying an external local gradient term, Pluim et al. (2000) incorporated both orientation and magnitude of gradient into the conventional mutual information measure to ensure the alignment of locations of tissue transitions in images. The plots of the similarity measure as a function of misalignments around ground truth indicate that the registration function of the combined measure is smoother than that of the standard mutual information measure.

On the other hand, some researchers applied mutual information or other entropic similarity measures on high-dimensional spatial feature extracted from the images, and adopted estimators for high-dimensional distributions (Ma et al., 2000; Neemuchwala et al., 2005; Sabuncu and Ramadge, 2003).

Region-based information: Rueckert et al. (2000) exploited second-order mutual information measure to include local spatial information present with the neighboring point pairs, by considering the co-occurrence of intensity pairs of adjacent points in individual images. This method is shown to be robust with respect to local intensity variation for non-rigid registration.

Moreover, Russakoff (2004) proposed the regional mutual information (RMI) measure by considering the co-occurrence of intensity groups of local neighborhood regions. To estimate the high-dimensional mutual information, they assumed that the high-dimensional joint distribution is approximately normally distributed. Various neighborhood radii were tested in their work, and it is demonstrated that RMI improves registration robustness.

Recently, Bardera et al. (2006) proposed to extend the normalized mutual information (Studholme et al., 1999) to consider the correspondence between voxel blocks. The estimation of the high-dimensional joint histogram was based on uniformly distributed random lines. Improved registration accuracy and robustness are demonstrated with the proposed method.

However, for the above adaptations, only the spatial information in a relatively local neighborhood is considered. In order to incorporate more global spatial information into the mutual information based registration framework, this paper defines and introduces a new spatial feature field, namely the maximum distance-gradient (MDG) vector field. The MDG feature encodes both the local edge information and spatial information at a global level which includes (1) the intensity difference, (2) the distance and (3) the direction of a voxel to a MDG source point. By incorporating the magnitude and orientation information of the MDG vector field, we propose an adaptation of the conventional mutual information based image registration framework. The overall similarity metric consists of two components: (1) the multi-dimensional mutual information measure on a newly formed two-element attribute vector space, and (2) an angle measure on the orientations of the MDG vector field of the images to be registered. To increase computational efficiency and robustness of the proposed method, we adopt a multi-resolution iterative registration process.

Based on the results using clinical three-dimensional CT and T1-weighted MRI image volumes from the retrospective image registration evaluation (RIRE) project (West et al., 1997), it is experimentally shown that, at different image resolutions, the proposed method can give longer capture ranges than the conventional mutual information based method and two of its adaptations (respectively proposed by Pluim et al. (2000) and Rueckert et al. (2000)) with incorporating spatial information. This can obviously make the multi-resolution image registration more robust. Moreover, the results of around 2000 randomized rigid registration experiments reveal that our method consistently gives much higher registration success rates than the aforementioned three related methods. Finally, it is demonstrated that our method can obtain high registration accuracy.

The organization of the paper is as follows. Section 2 formulates spatial information as a novel MDG vector field. Our objective function for multimodal image registration method is proposed in Section 3. Implementation details for the computation of the MDG vector field and also for the registration process are given in Section 4. Section 5 presents the experimental results and discussions. The conclusion is drawn in Section 6.

2. Maximum distance-gradient vector field

It is known that the conventional mutual information based registration method does not consider spatial information in the images. The aim of our approach is to incorporate spatial information into the mutual information measure in order to improve registration performance, particularly with respect to registration robustness. We define below a new spatial feature field, namely the maximum distance-gradient (MDG) vector field, to encode both the local edge information and globally defined spatial information, which includes (1) the intensity difference, (2) the distance, and (3) the direction of a voxel to a MDG source point. The adaptation of the mutual information measure by associating the MDG vector field with the conventional voxel intensity is proposed in Section 3.

2.1. Conventional image gradient

Image gradient is a commonly adopted representation of spatial information. The conventional image gradient is a first derivative based operator, that is to say,

$$\nabla I(\mathbf{v}) = \frac{\partial I(\mathbf{v})}{\partial x}\vec{i}_x + \frac{\partial I(\mathbf{v})}{\partial y}\vec{i}_y + \frac{\partial I(\mathbf{v})}{\partial z}\vec{i}_z,$$

where $\mathscr{I} = \{I(\mathbf{v})\}$ is an intensity image volume, $\mathbf{v} = (x, y, z)$ denotes a voxel position, and $\vec{i}_x$, $\vec{i}_y$ and $\vec{i}_z$ are unit vectors in the principal directions, respectively. At each point in the image, the gradient vector points in the direction of the largest possible intensity increase, and the length of the gradient vector corresponds to the rate of change along that direction.

Fundamentally, the conventional image gradient is locally defined and normally used to detect the significance of object boundaries, where voxels change their intensity suddenly. By computing the magnitude of the image gradient, voxels at object boundaries would give large values, while voxels within homogeneous regions would give small and almost constant values. Specifically in medical images, locations around tissue transitions, which only occupy a very small proportion of the whole image volume, present high information value (i.e., large gradient magnitude). On the other hand, a large amount of voxels (i.e., locations within background regions and anatomical structures) have small and almost constant gradient magnitude and thus the presented information for image registration is very limited. Consequently, this may limit the use and effect of image gradient features in voxel-based image registration. A graphical illustration for these behaviors is shown in Fig. 6 and a further discussion is given in Section 2.2.2.

2.2. Maximum distance-gradient (MDG)

In this subsection, we define a new spatial feature field, which we call the maximum distance-gradient (MDG) vector field. It contains not only the local edge information, but also spatial information at a global level, which encodes (1) the intensity difference, (2) the distance and (3) the direction of a voxel to a MDG source point, which will be defined in the next subsection. Moreover, the magnitude of the MDG feature varies smoothly and gradually from the object boundaries towards the interiors of homogeneous image regions.

2.2.1. Definition

To define the MDG vector field, we begin by defining a *distance-gradient* operator, ∇_d, on two arbitrary voxels in an image. The distance-gradient of two voxels, $\mathbf{v}_1$ and $\mathbf{v}_2$, in an image volume $\mathscr{I} = \{I(\mathbf{v})\}$ is a three-dimensional vector (or a two-dimensional vector when $\mathbf{v} = (x, y)$), which is defined as

$$\nabla_d I(\mathbf{v}_1, \mathbf{v}_2) = (I(\mathbf{v}_1) - I(\mathbf{v}_2))\frac{\mathbf{v}_1 - \mathbf{v}_2}{|\mathbf{v}_1 - \mathbf{v}_2|^2}$$

$$= \frac{I(\mathbf{v}_1) - I(\mathbf{v}_2)}{|\mathbf{v}_1 - \mathbf{v}_2|}\vec{n}_{\mathbf{v}_1\mathbf{v}_2}, \qquad (1)$$

with

$$\vec{n}_{\mathbf{v}_1\mathbf{v}_2} = \frac{\mathbf{v}_1 - \mathbf{v}_2}{|\mathbf{v}_1 - \mathbf{v}_2|}.$$

Here $|\cdot|$ denotes vector magnitude. Specifically, vector $\nabla_d I(\mathbf{v}_1, \mathbf{v}_2)$ points in the direction of intensity increase between $\mathbf{v}_1$ and $\mathbf{v}_2$, i.e., from $\mathbf{v}_2$ to $\mathbf{v}_1$ if $I(\mathbf{v}_1) > I(\mathbf{v}_2)$, and vice versa. On the other hand, the vector length (or magnitude) is in direct ratio to the intensity difference but in inverse ratio to the location distance of $\mathbf{v}_1$ and $\mathbf{v}_2$. This means that $|\nabla_d I(\mathbf{v}_1, \mathbf{v}_2)|$ measures the average rate at which the intensity function $I(\mathbf{v})$ changes between $\mathbf{v}_1$ and $\mathbf{v}_2$. Note that Eq. (1) is valid for any two voxels as long as they are from different locations.

With the above definition, a MDG vector field, $\mathscr{G} = \{G(\mathbf{v})\}$, of the image can be derived by using

$$G(\mathbf{v}) = \nabla_d I(\bar{\mathbf{v}}, \mathbf{v}), \qquad (2)$$

where

$$\bar{\mathbf{v}} = \arg\max_{\mathbf{v}' \in \Omega} |\nabla_d I(\mathbf{v}', \mathbf{v})|. \qquad (3)$$

Here Ω is the image volume domain. In other words, for any voxel $\mathbf{v}$, there exists a corresponding voxel, $\bar{\mathbf{v}}$, which gives the maximum distance-gradient magnitude with $\mathbf{v}$ among all voxels in the image volume. Then, the MDG vector at $\mathbf{v}$ is defined as the distance-gradient between $\bar{\mathbf{v}}$ and $\mathbf{v}$. Hereafter we call $\bar{\mathbf{v}}$ the MDG source point of $\mathbf{v}$.

According to this formulation, the magnitude of $G(\mathbf{v})$ gives the maximum average rate at which the intensity function I changes between $\mathbf{v}$ and any other voxels in the image volume. Its orientation is in the direction of intensity increase between $\mathbf{v}$ and $\bar{\mathbf{v}}$. Based on Eqs. (1) and (2), $\bar{\mathbf{v}}$ (the MDG source point of $\mathbf{v}$) is generally located at or around an object boundary, in order to simultaneously maintain relatively large intensity difference and relatively small location distance. Moreover, for empirical justification, the spatial distributions of the MDG source points in a two-dimensional slice extracted from the three-dimensional image volume are plotted for both CT and T1-weighted MR images, as shown in Figs. 1a and b, respectively. Note
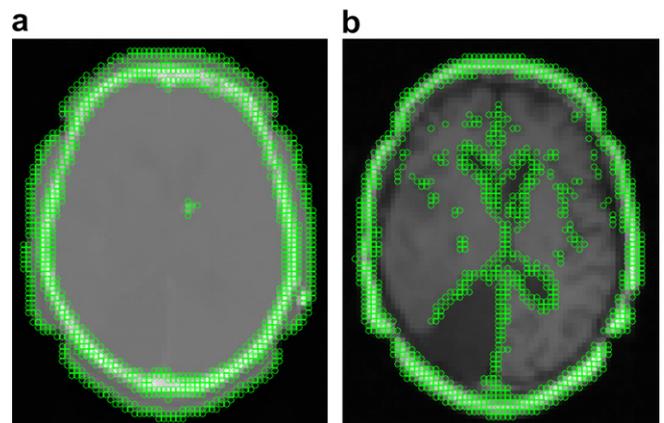

Fig. 1. The spatial distributions of the MDG source points for (a) CT and (b) MR images. Note that, although two-dimensional slices are shown, the computations of MDG source points were done in three dimensions. The MDG source points are marked with green circles. From the figure, it is observed that most of the MDG source points are located at or around the object boundary in both images.
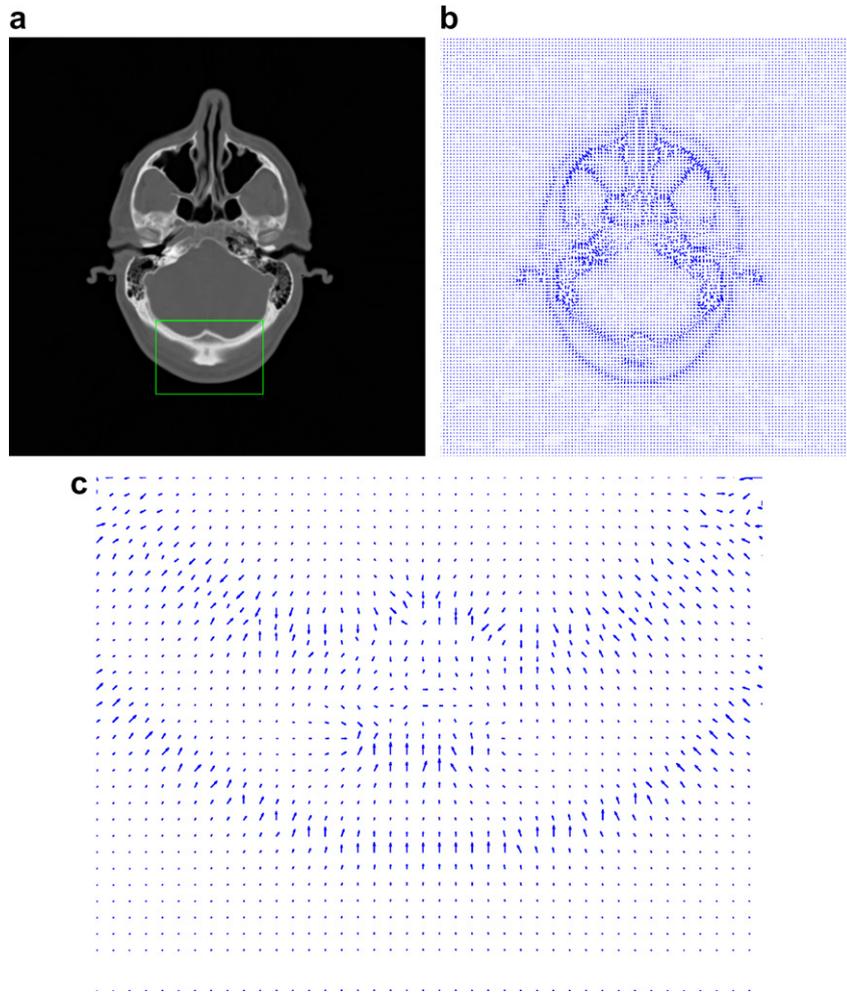
Fig. 2. (a) An in-plane two-dimensional CT slice. (b) The signed-MDG feature field of (a). (c) A close-up of the field for the ROI within the box in (a). The computation is just performed in the in-plane two-dimensional image slice for a better representation of vectors.

that the computations of MDG sources were done in three dimensions. The MDG source points are marked with green circles[1]. From the figure, it is observed that most of the MDG source points are located at or around the object boundary in both images.

More examples of the MDG vector field computed for real clinical CT and T1-weighted MR images are shown in Figs. 2 and 3. For a clear representation of vectors, in-plane two-dimensional slices, as shown in Figs. 2a and 3a, are extracted from three-dimensional clinical CT and MR image volumes obtained from the RIRE project. Then, the computed MDG vector fields for these in-plane slices are shown in Figs. 2b and 3b (see Section 4 for the implementation details of the MDG vector field computation). Note that, in order to delineate the locations of the object boundaries, Figs. 2b and 3b plot the signed-MDG feature field, which will be defined in Eq. (8),

$$\text{sign}(I(\bar{\mathbf{v}}) - I(\mathbf{v}))G(\mathbf{v}),$$

where the function $\text{sign}(\cdot)$ indicates the sign of the intensity difference. That is to say, instead of pointing in the direction of intensity increase, a vector always points in the direction from a point to its MDG source point. Figs. 2c and 3c show close-ups of the MDG vector fields for the regions of interest (ROI) selected by boxes as shown in Figs. 2a and 3a.

2.2.2. Properties

Compared with the conventional image gradient, the MDG vector field of the CT and MR image volumes has at least two superior characteristic features, Feature I and Feature II:

Feature I. When a voxel $\mathbf{v}$ is at or very close to an object boundary, its MDG vector $G(\mathbf{v})$ generally approximates the conventional image gradient $\nabla I(\mathbf{v})$, with respect to both orientation and magnitude.

It can be illustrated by using a clinical CT image volume (Fig. 4a) and a clinical T1-weighted MR image volume (Fig. 4f) obtained from the RIRE project. For the CT image volume, the gradient and MDG vector fields were

---

[1] For interpretation of the references to colour in the text, the reader is referred to the web version of this article.
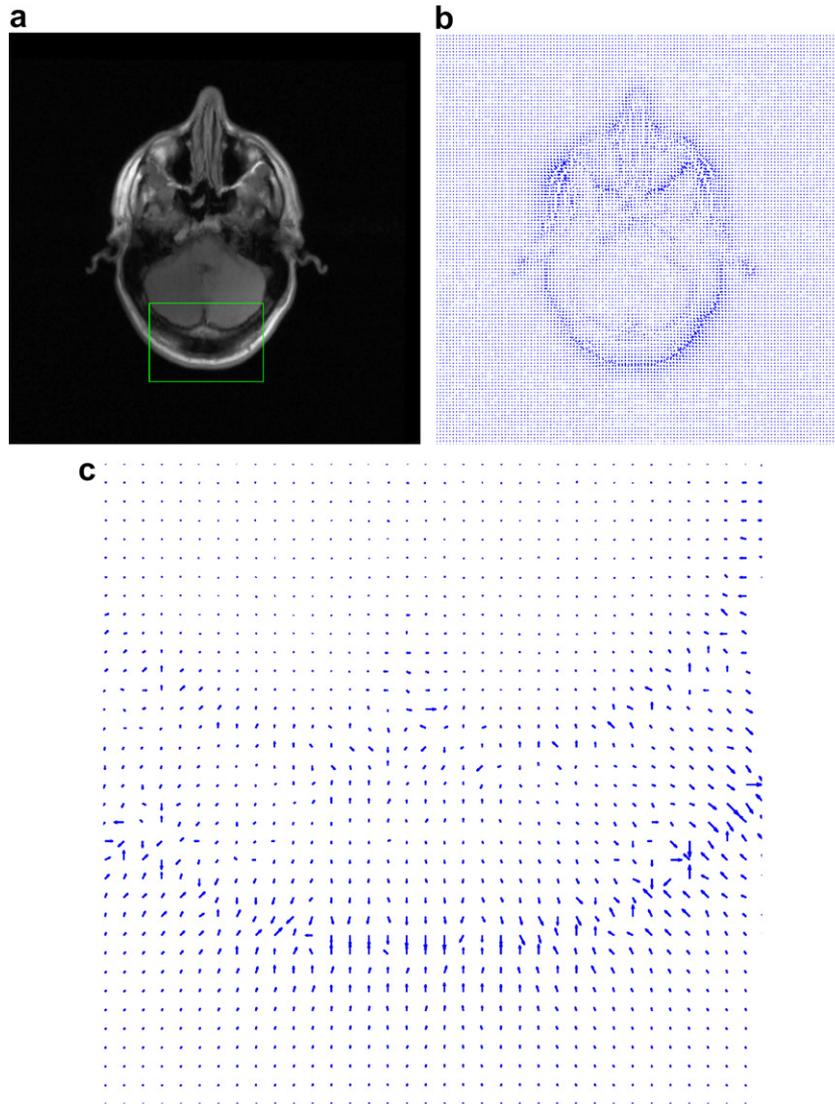
Fig. 3. (a) An in-plane two-dimensional MR slice. (b) The signed-MDG feature field of (a). (c) A close-up of the field for the ROI within the box in (a). The computation is just performed in the in-plane two-dimensional image slice for a better representation of vectors.

estimated, and Figs. 4b and 4c show the estimated gradient and MDG magnitudes, respectively. Boundary points were selected based on a threshold, which was set to 33% of the maximum gradient magnitude. Points (or voxels) having gradient magnitude above the threshold were marked as the boundary points. At each boundary point, the corresponding gradient and MDG vectors were kept (see Fig. 4d for gradient and Fig. 4e for MDG); otherwise, vectors were set to zero at the non-boundary points. For all the marked boundary points, pairs of gradient and MDG magnitudes are plotted in Fig. 5a and the cross-correlation is 0.9648. Similarly, all pairs of gradient and MDG orientations are plotted in Fig. 5b and the cross-correlation is 0.9522.

For the MR image volume, Figs. 4g and h show the estimated gradient and MDG magnitudes, respectively. Using 33% of the maximum gradient magnitude as the threshold, points (or voxels) having gradient magnitude above the

threshold were marked as the boundary points. The gradient and MDG magnitudes on the boundary points are plotted in Figs. 4i and j, respectively. For all the boundary points, pairs of gradient and MDG magnitudes are plotted in Fig. 5c and the cross-correlation is 0.9427. For orientations, the cross-correlation is 0.9518, and pairs of gradient and MDG orientations are plotted in Fig. 5d. In the above CT image volume and MR image volume, between gradient and MDG vectors, it is observed that the values of magnitude and orientation cross-correlations are relatively high at the marked boundary points.

This feature can also be partly explained as follows. When a voxel $\mathbf{v}$ is at or very close to an object boundary, if its MDG source point $\bar{\mathbf{v}}$ is very close to $\mathbf{v}$, then the vector $\Delta \mathbf{v} = \bar{\mathbf{v}} - \mathbf{v}$ becomes very small. Based on the Taylor's theorem (Spivak, 1980), we have

$$I(\bar{\mathbf{v}}) = I(\mathbf{v} + \Delta \mathbf{v}) \approx I(\mathbf{v}) + \nabla I(\mathbf{v}) \cdot \Delta \mathbf{v}.$$
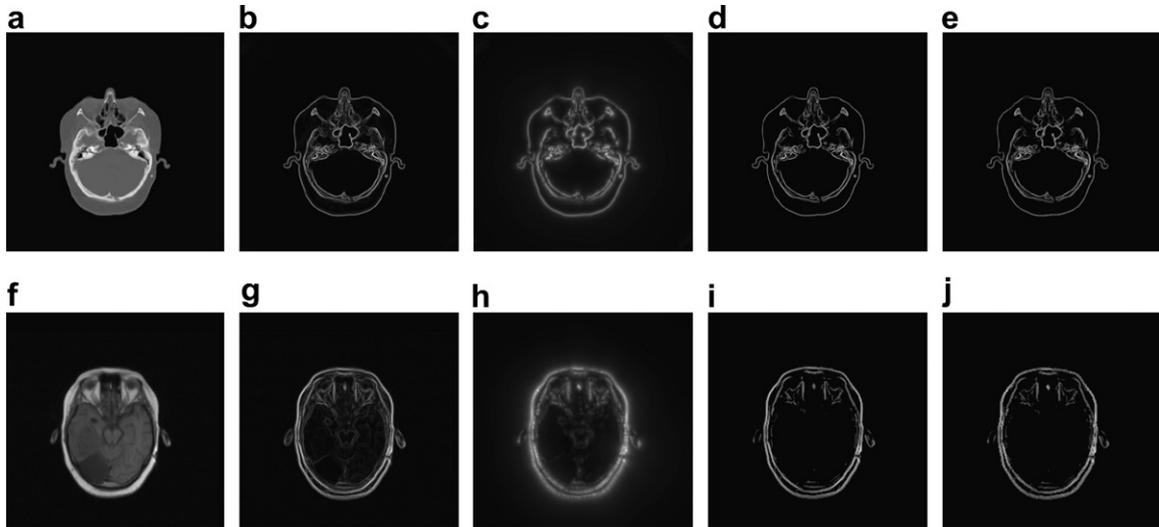
Fig. 4. (a) An in-plane two-dimensional slice selected from a clinical CT image volume. (b) and (c) The corresponding slices for the gradient magnitude and the MDG magnitude, respectively. (d) The corresponding slice which only shows the values of gradient magnitude, obtained from (b), at the boundary points. (e) The corresponding slice which only shows the values of MDG magnitude, obtained from (c), at the boundary points. (f) An in-plane two-dimensional slice selected from a clinical MR image volume. (g) and (h) The corresponding slices for the gradient magnitude and the MDG magnitude, respectively. (i) The corresponding slice which only shows the values of gradient magnitude, obtained from (g), at the boundary points. (j) The corresponding slice which only shows the values of MDG magnitude, obtained from (h), at the boundary points. All computations are performed in the three-dimensional image volume.
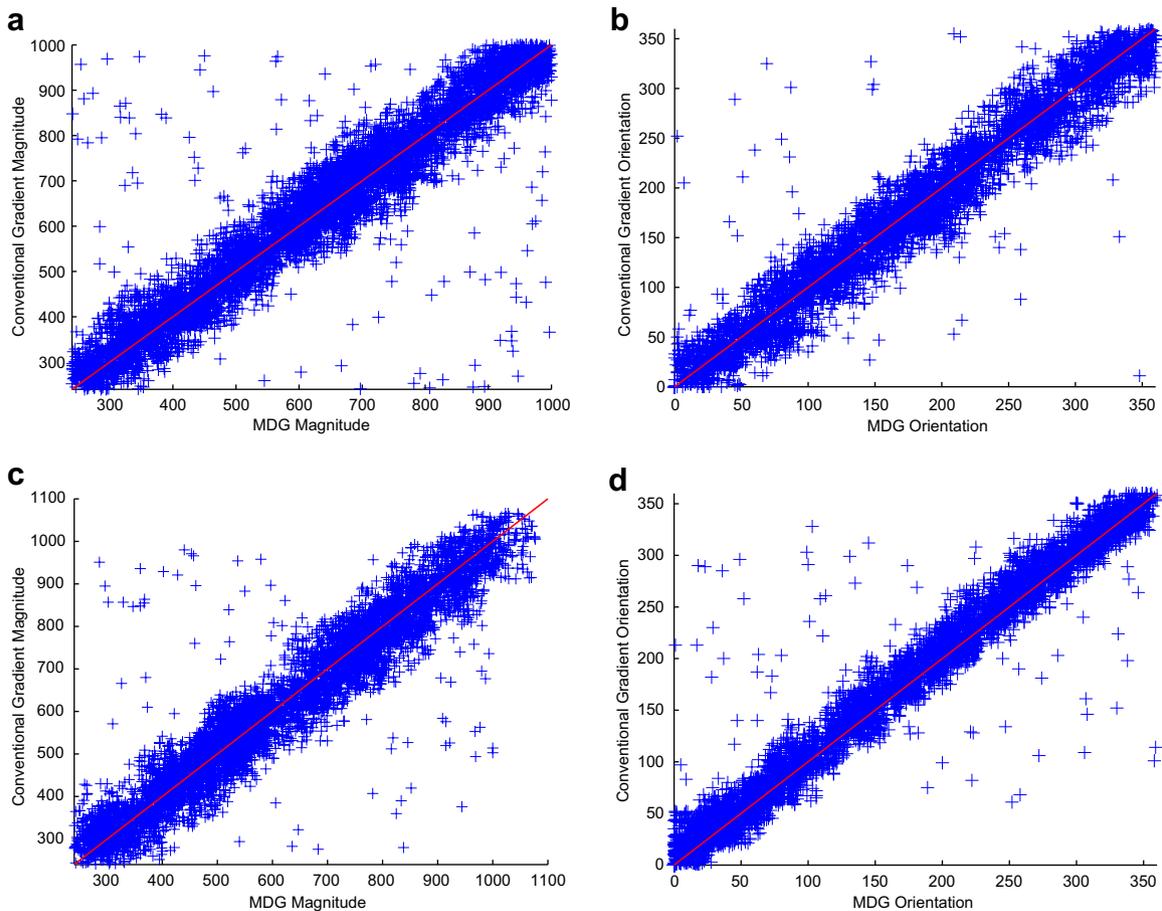

Fig. 5. (a) and (c) Scatter plots which show all pairs of gradient and MDG magnitudes at the boundary points in the clinical CT and MR image volumes, respectively. (b) and (d) Scatter plots which show all pairs of gradient and MDG orientations at the boundary points in the clinical CT and MR image volumes, respectively.

Subsequently, based on Eq. (2), it is true that

$$G(\mathbf{v}) = \nabla_d I(\bar{\mathbf{v}}, \mathbf{v}) = \nabla_d I(\mathbf{v} + \Delta\mathbf{v}, \mathbf{v}) \approx \frac{\nabla I(\mathbf{v}) \cdot \Delta\mathbf{v}}{|\Delta\mathbf{v}|} \frac{\Delta\mathbf{v}}{|\Delta\mathbf{v}|}$$

$$= |\nabla I(\mathbf{v})| \cos\hat{\theta} \frac{\Delta\mathbf{v}}{|\Delta\mathbf{v}|}, \tag{4}$$

where $\hat{\theta}$ is the angle between $\nabla I(\mathbf{v})$ and $\Delta\mathbf{v}$.

On the other hand, since it is assumed that $\bar{\mathbf{v}}$ is very close to $\mathbf{v}$, the MDG magnitude for $\mathbf{v}$ can be found by searching the voxels close to $\mathbf{v}$. In this situation, according to the definition of the MDG vector field (Eqs. (2), (3)), the MDG magnitude for $\mathbf{v}$ is given by

$$|G(\mathbf{v})| = \max_{\mathbf{h}} |\nabla_d I(\mathbf{v} + \mathbf{h}, \mathbf{v})|, \tag{5}$$

where $\mathbf{h}$ denotes a small vector, $\mathbf{v} + \mathbf{h} \in \Omega$, and Ω represents the image volume domain. In other words, we just need to search voxels $\mathbf{v} + \mathbf{h}$ which are close to $\mathbf{v}$. When $|\mathbf{h}|$ becomes very small, using Eq. (4) and Eq. (5) can be approximated by

$$|G(\mathbf{v})| \approx \max_{\theta} |\nabla I(\mathbf{v})||\cos\theta| = |\nabla I(\mathbf{v})|, \tag{6}$$

where θ is the angle between $\nabla I(\mathbf{v})$ and $\mathbf{h}$. This shows that the magnitude of $G(\mathbf{v})$ approximates that of $\nabla I(\mathbf{v})$, when $\mathbf{v}$ is located at or very close to an object boundary and $\bar{\mathbf{v}}$ is very close to $\mathbf{v}$. Furthermore, by combining Eqs. (4) and (6), we derive $\hat{\theta} \approx 0$, which implies that the orientation of $G(\mathbf{v})$ tends to coincide with that of $\nabla I(\mathbf{v})$ in this situation.

Feature II. The MDG vector field of homogeneous regions, e.g. anatomical structures and background regions, provides rich information related to the intensity change, the distance, and the direction of a voxel to a certain object boundary point, which is the MDG source point. This property differs markedly from the limited information provided by the conventional image gradient in the homogeneous regions.

Specifically, when the voxel position varies from the object boundaries towards the interiors of homogeneous regions, the magnitude of the MDG vector *decays smoothly and gradually*. It is worth pointing out that the MDG source points of the neighboring locations may not be the identical or neighboring points. Moreover, when the voxels are located relatively far away from the boundaries, such a decrement in magnitude can be very small.

Below we give a mathematical analysis on the magnitude variation of the MDG vectors within the homogeneous image regions. Suppose $\mathbf{v}_a$ and $\mathbf{v}_b$ are two neighboring locations in a homogeneous region, and $\bar{\mathbf{v}}_a$ and $\bar{\mathbf{v}}_b$ are their MDG source points, respectively. To estimate their magnitude difference, $\||G(\mathbf{v}_a)| - |G(\mathbf{v}_b)|\|$, we first compute

$$\||\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_a)| - |\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_b)|\|$$

$$= \left| \frac{|I(\bar{\mathbf{v}}_a) - I(\mathbf{v}_a)|}{|\bar{\mathbf{v}}_a - \mathbf{v}_a|} - \frac{|I(\bar{\mathbf{v}}_a) - I(\mathbf{v}_b)|}{|\bar{\mathbf{v}}_a - \mathbf{v}_b|} \right|. \tag{7}$$

Since $\mathbf{v}_a$ and $\mathbf{v}_b$ are very close (i.e., adjacency), and also $I(\mathbf{v}_a)$ and $I(\mathbf{v}_b)$ are very similar (i.e., homogeneity), we have

$$|I(\bar{\mathbf{v}}_a) - I(\mathbf{v}_b)| = |I(\bar{\mathbf{v}}_a) - I(\mathbf{v}_a)| + \varepsilon_I,$$

and

$$|\bar{\mathbf{v}}_a - \mathbf{v}_b| = |\bar{\mathbf{v}}_a - \mathbf{v}_a| + \varepsilon_{\mathbf{v}},$$

where ε_I and $\varepsilon_{\mathbf{v}}$ are two small values, which approach 0. Let $\Delta\mathbf{v}_a = |\bar{\mathbf{v}}_a - \mathbf{v}_a|$ and $\Delta I_a = |I(\bar{\mathbf{v}}_a) - I(\mathbf{v}_a)|$. Since $\mathbf{v}_a$ and $\mathbf{v}_b$ are away from the boundaries, it is reasonable to assume that $\Delta\mathbf{v}_a > 2|\varepsilon_{\mathbf{v}}|$. Consequently, Eq. (7) becomes

$$\||\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_a)| - |\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_b)|\| = \left| \frac{\Delta I_a}{\Delta\mathbf{v}_a} - \frac{\Delta I_a + \varepsilon_I}{\Delta\mathbf{v}_a + \varepsilon_{\mathbf{v}}} \right|$$

$$= \left| \frac{\Delta I_a \cdot \varepsilon_{\mathbf{v}}}{\Delta\mathbf{v}_a(\Delta\mathbf{v}_a + \varepsilon_{\mathbf{v}})} - \frac{\varepsilon_I}{\Delta\mathbf{v}_a + \varepsilon_{\mathbf{v}}} \right|$$

$$\leqslant \frac{\Delta I_a}{\Delta\mathbf{v}_a} \cdot \frac{2|\varepsilon_{\mathbf{v}}|}{\Delta\mathbf{v}_a} + \frac{2|\varepsilon_I|}{\Delta\mathbf{v}_a}.$$

It is obvious that, as long as $\mathbf{v}_a$ is not close to an object boundary, we can find a small value ε to satisfy

$$\||\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_a)| - |\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_b)|\| < \varepsilon,$$

and then with similar computation,

$$\||\nabla_d I(\bar{\mathbf{v}}_b, \mathbf{v}_b)| - |\nabla_d I(\bar{\mathbf{v}}_b, \mathbf{v}_a)|\| < \varepsilon.$$

Based on Eqs. (2) and (3), we have

$$|G(\mathbf{v}_a)| - |G(\mathbf{v}_b)| > |\nabla_d I(\bar{\mathbf{v}}_b, \mathbf{v}_a)| - |\nabla_d I(\bar{\mathbf{v}}_b, \mathbf{v}_b)|$$
$$> -\varepsilon,$$

and

$$|G(\mathbf{v}_a)| - |G(\mathbf{v}_b)| < |\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_a)| - |\nabla_d I(\bar{\mathbf{v}}_a, \mathbf{v}_b)| < \varepsilon.$$

The above two equations represent that the magnitude variation of these two MDG vectors can be bounded by a small value, that is

$$\||G(\mathbf{v}_a)| - |G(\mathbf{v}_b)|\| < \varepsilon.$$

As such, it can be verified that the magnitude of the MDG vector within homogeneous regions changes smoothly and gradually.

To summarize, the MDG vector field can encode the local edge information as well as spatial information in a more global sense which includes (1) the intensity difference, (2) the distance, and (3) the direction of a voxel to a MDG source point.

To give a comparative example, the magnitude of the conventional image gradient and that of the MDG vector field of a three-dimensional clinical CT volume obtained from the RIRE project are computed. All computations are performed in the three-dimensional image volume. A slice extracted from the volume is shown in Fig. 6a, while Figs. 6b and c, respectively present the corresponding slices for the magnitude of gradient and that of the MDG vector.

Note that values from individual images are re-scaled to [−0.5, 0.5] for fair comparison.

It is observed that the gradient magnitude (Fig. 6b) only exhibits the sharp edge information, which is also similarly presented by the magnitude of the MDG vector (Fig. 6c). In addition, unlike Fig. 6b, the smooth and gradual change in the magnitude of the MDG vector can be demonstrated
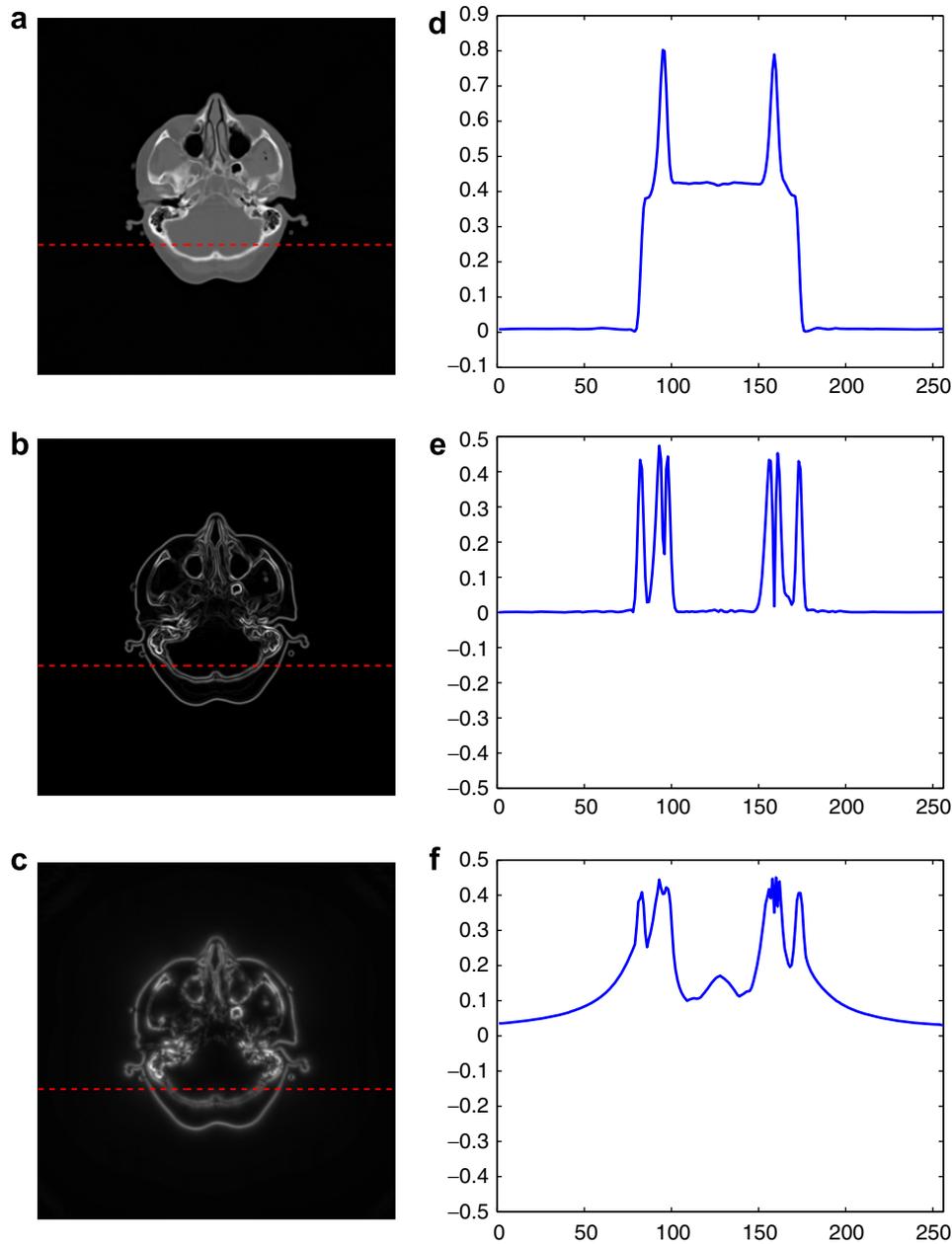
Fig. 6. (a) is a slice selected from a clinical CT image volume. (b) and (c) are the corresponding slices for the gradient magnitude and the MDG magnitude. (d)-(f) Are value profiles of lines in (a)–(c), which are marked as dashed lines. All computations are performed in the three-dimensional image volume.

in Fig. 6c, for instance, in regions close to the boundaries. However, due to the limitation of image quality in the paper, smooth magnitude variation within the background regions and anatomical structures may not be clearly displayed in Fig. 6c.

For a detailed description, Figs. 6d–f, respectively present the value profiles of the same line (marked as dashed lines) in Figs. 6a–c. As shown in the figures, feature values in Fig. 6e are very sparse, in which the overwhelming majority are small and constant. In contrast, the value variation from the object boundaries towards the interiors of homogeneous regions in Fig. 6f is smooth and gradual. It is worth noting that, although there is little intensity varia-

tion in the middle of the line in Fig. 6d, an evident and smooth saddle can be found in Fig. 6f located at the corresponding position. The raised white boundary slightly below the line causes this saddle. This is because, unlike the gradient operator, the MDG vector field is defined at a more global level.

2.3. Signed-MDG

In order to distinguish between voxels of objects with different intensities based on the magnitude of the MDG vector, we further introduce a so-called signed-MDG feature field, $\hat{\mathscr{G}} = \{\widehat{G}(\mathbf{v})\}$, as follows:

$$\|\widehat{G}(\mathbf{v})\| := \text{sign}(I(\bar{\mathbf{v}}) - I(\mathbf{v}))|G(\mathbf{v})|,$$
$$\text{Orientation of } \widehat{G}(\mathbf{v}) := \text{Orientation of } G(\mathbf{v}), \tag{8}$$

where $\|\widehat{G}(\mathbf{v})\|$ denotes the magnitude of $\widehat{G}(\mathbf{v})$, $\bar{\mathbf{v}}$ is the MDG source of $\mathbf{v}$, and the function sign($\cdot$) indicates the sign of the intensity difference.

According to this modified definition, the magnitude of a signed-MDG feature can be either positive or negative with respect to the intensity difference between a voxel and its MDG source, i.e., a voxel of relatively low intensity would have a positive signed-MDG magnitude and vice versa. Note that the orientation as well as the absolute value of the magnitude of $\widehat{G}(\mathbf{v})$ are identical with those of $G(\mathbf{v})$. Hereafter, the absolute value of magnitude of $\widehat{G}(\mathbf{v})$ is referred to as absolute magnitude of $\widehat{G}(\mathbf{v})$.

For comparison, Figs. 7a and b, respectively show the corresponding slice and value profile of the magnitude of the obtained signed-MDG for the aforementioned CT image volume as shown in Fig. 6. The magnitude of signed-MDG preserves all the properties presented by the magnitude of the MDG vector (i.e., Fig. 6c). Furthermore, as shown in Fig. 7b, the voxels of objects with different intensities are distinguishable by the sign of the magnitude. Hereafter, we adopt the signed-MDG field to encode spatial information for registration tasks. The implementation
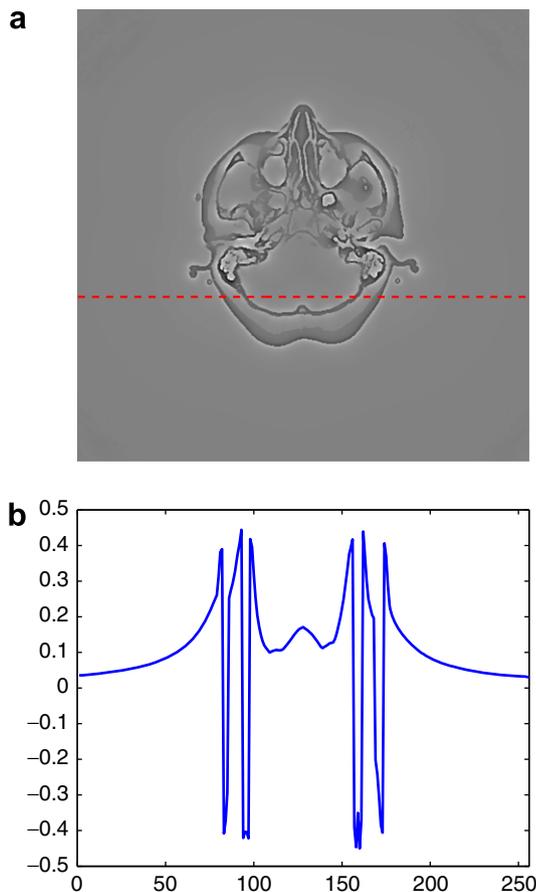

Fig. 7. (a) Corresponding slice of the magnitude of the signed-MDG for the aforementioned CT slice shown in Fig. 6a. (b) The value profile of the line in (a), which is marked as the dashed line.

details for the computation of signed-MDG field is given in Section 4.1.

3. Incorporation of spatial information in image registration

As we have discussed in Section 2, the signed-MDG feature field encodes both local edge information and spatial information at a global level:

- The orientation of the signed-MDG feature is along the direction of a voxel and its MDG source, which is located at a MDG source point.
- The absolute magnitude of the signed-MDG measures the maximum average rate at which the intensity function changes between a voxel and any other voxels in the image volume. The sign of the magnitude indicates the relative intensity level between a voxel and its MDG source point.

In Sections 3.1 and 3.2, we separately incorporate the two components of the signed-MDG feature field (i.e., magnitude and orientation) into the registration process, and propose a new objective function in Section 3.3.

3.1. Multi-dimensional mutual information

To make use of the magnitude of the signed-MDG feature field, we associate it with the voxel intensity to form a two-dimensional feature space, $\mathscr{Z} = \{\mathbf{Z}(\mathbf{v})\}$. That is to say, for a voxel $\mathbf{v}$, it has a two-element attribute vector $\mathbf{Z}(\mathbf{v}) = (I(\mathbf{v}), \|\widehat{G}(\mathbf{v})\|)$. The new feature space $\mathscr{Z}$ gives at least one advantage over the intensity space. It differentiates voxels more subtly than the intensity space alone, especially for voxels within homogeneous image regions which generally occupy a large proportion of the whole image volume. In fact, voxels with an identical intensity would have different magnitudes of the signed-MDG because they may have either different MDG source points or different distances from their MDG source points.

The *first* element in our overall objective function is a measure of the similarity between the newly defined two-dimensional feature space $\mathscr{Z}$ of the two images to be registered. Mutual information, which is adopted in our work, has been extensively exploited as a similarity measure on random variables. It is an information-theoretic measure, and does not make limiting assumptions on the relationship (for instance, linear relationship) between variables. Mutual information was originally introduced as a similarity measure on intensity space for medical image registration, and correct (or optimal) registration is meant to maximize the mutual information of intensity spaces of two images over the registration transformation space.

The mutual information MI of two random variables A and B is defined in terms of entropy in the following way (Cover and Thomas, 1991; Vajda, 1989)

$$\text{MI}(A, B) = H(A) + H(B) - H(A, B).$$

Here $H(\cdot)$ is the Shannon entropy of a random variable with probability distribution p, and is defined as $-\sum p \log p$. $H(\cdot, \cdot)$ is the joint Shannon's entropy of two random variables with joint probability distribution $p(\cdot, \cdot)$, and is defined as $-\sum \sum p(\cdot, \cdot) \log p(\cdot, \cdot)$.

Let r and f denote the *reference* and *floating* images to be registered, respectively. Then, the mutual information measure on the new attribute vector space $\mathscr{Z}$ of two images with a registration transformation $\mathbf{T}$ is defined as

$$\mathrm{MI}(\mathbf{T}) = MI_{\mathbf{T}}(\mathscr{Z}_f, \mathscr{Z}_r)$$
$$= H_{\mathbf{T}}(\mathscr{Z}_f) + H_{\mathbf{T}}(\mathscr{Z}_r) - H_{\mathbf{T}}(\mathscr{Z}_f, \mathscr{Z}_r).$$

Since $\mathbf{Z}(\mathbf{v})$ is two-dimensional, the three terms in the equation are defined as

$$H_{\mathbf{T}}(\mathscr{Z}_f) = -\sum p_{\mathbf{T}}(I_f, \|\widehat{G}_f\|) \log p_{\mathbf{T}}(I_f, \|\widehat{G}_f\|),$$
$$H_{\mathbf{T}}(\mathscr{Z}_r) = -\sum p_{\mathbf{T}}(I_r, \|\widehat{G}_r\|) \log p_{\mathbf{T}}(I_r, \|\widehat{G}_r\|),$$
$$H_{\mathbf{T}}(\mathscr{Z}_r, \mathscr{Z}_f) = -\sum p_{\mathbf{T}}(I_r, \|\widehat{G}_r\|, I_f, \|\widehat{G}_f\|)$$
$$\log p_{\mathbf{T}}(I_r, \|\widehat{G}_r\|, I_f, \|\widehat{G}_f\|).$$

The computation of these joint probability distributions is given in Section 4.2.1.

3.2. Contribution from orientation of signed-MDG

The *second* element in our objective function is the contribution from the orientation information of the signed-MDG feature field. The essential idea is an assumption, which is always valid in practice according to our empirical experience: for two precisely aligned multi-modal images, the majority of the corresponding voxel locations are very likely to produce signed-MDG features with similar orientations. The directions can be either almost identical or opposite.

This is because the signed-MDG feature always points along the direction from a voxel to a MDG source point, which presents the maximum average rate of intensity change with that voxel. Meanwhile, two images acquired by different imaging modalities (e.g. CT and MR-T1, etc.) fundamentally depict the same anatomical structures. As a result, the signed-MDG features of the corresponding voxel locations from two images could present relatively coincident orientations in general.

In our method, we quantitatively measure the orientation difference between the signed-MDG features from two images at voxel location $\mathbf{v}$ as

$$|\cos(\theta(\mathbf{v}))|.$$

Here $\theta(\mathbf{v})$ is the angle between two signed-MDG features and is obtained by

$$\theta(\mathbf{v}) = \arccos \frac{G_r(\mathbf{v}) \cdot G_f(\mathbf{v})}{|G_r(\mathbf{v})||G_f(\mathbf{v})|}.$$

There may be a concern about the aforementioned assumption: it is known that multi-modal images do not necessar-

ily delineate the same tissue transitions, which implies that boundaries that appear in a certain modality may be absent or less noticeable in another modality. However, based on our empirical experience, this assumption is generally valid for the voxel locations from the common structures present in both images, particularly for those from background regions, which occupy a very large proportion of the whole image volume. For instance, more than 70% of the total voxels belong to background for the datasets obtained from the Retrospective Image Registration Evaluation (RIRE) Project.

For illustration, the angle measure is computed on the signed-MDG fields of two accurately registered three-dimensional CT and MR-T1 image volumes. We then plot the relative frequency of the obtained angle measure values for the overall three-dimensional image volume in Fig. 8. It is observed that the angle measure produces values of $> 0.866 (= \cos 30°)$ at more than 70% of total voxels. Note that here the signed-MDG features are three-dimensional as the computations are performed on the overall image volumes.

On the other hand, unlike the optimal registration discussed above, a misalignment would dramatically reduce the relative coincidence in orientation, and result in large angles between the signed-MDG features of corresponding voxels. For instance, after an in-plane translation of the above mentioned MR-T1 image volume, the relative frequency of the obtained values of angle measure becomes much more dispersed, as shown in Fig. 9. It is observed that only around 40% of total voxels present values of $> 0.866 (= \cos 30°)$ for the angle measure.

In order to get the overall contribution from the orientation information of the signed-MDG field for registration tasks, the summation of the angle measure for all voxels in the volume of overlap of the reference image r and the floating image f is adopted. As such, at a registration transformation $\mathbf{T}$, we have
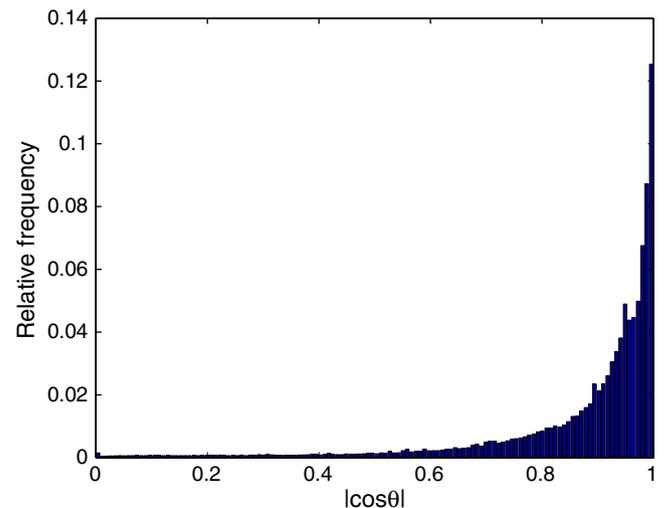

Fig. 8. The relative frequency of the obtained values of angle measure for the overall three-dimensional volume.
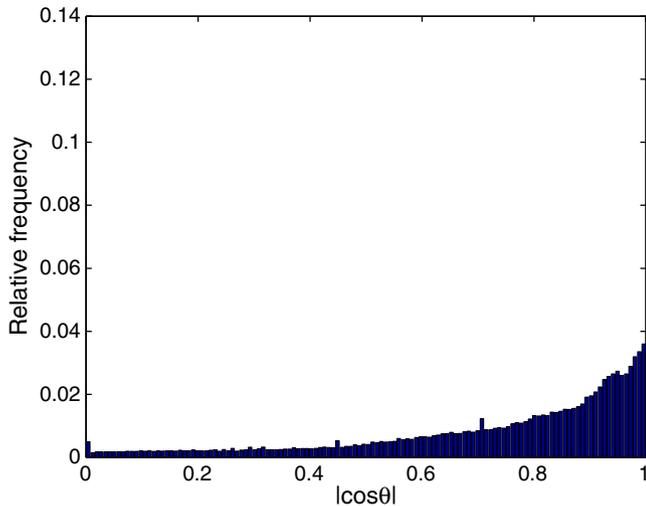
Fig. 9. The relative frequency of the obtained values of angle measure for the same image volume pair, as mentioned in Fig. 8, but with an in-plane misregistration.

$$D(\mathbf{T}) = \sum_{\mathbf{v} \in (f \cap r)} |\cos(\theta_{\mathbf{T}}(\mathbf{v}))|$$

as the second component in our registration objective function.

3.3. Registration objective function

In order to combine the two elements proposed in the above two subsections, we adopt multiplication rather than addition to reduce the need for normalization. As a consequence, at a registration transformation $\mathbf{T}$, the total similarity metric is defined as

$$S(\mathbf{T}) = MI(\mathbf{T}) \cdot D(\mathbf{T}). \tag{9}$$

As discussed above, both measures, $MI(\mathbf{T})$ and $D(\mathbf{T})$, achieve the maximal values at the accurate (or optimal) registration. Therefore, the overall registration approach is to estimate the optimal transformation $\widehat{\mathbf{T}}$ by using

$$\widehat{\mathbf{T}} = \arg\max_{\mathbf{T}} S(\mathbf{T}). \tag{10}$$

With this objective function, both the magnitude and orientation information of the signed-MDG feature field are included into the registration process.

4. Implementation

In this section, the implementation details for the computation of the signed-MDG feature field and also for our proposed multi-resolution image registration method are presented.

4.1. Signed-MDG

In our work, the signed-MDG field is computed by separating it into positive and negative components. Then, the two components are calculated based on the Voronoi diagram (Preparata and Shamos, 1985) by sequentially pro-

cessing voxels in intensity-decreasing and intensity-increasing orders, respectively.

During either procedure, we keep updating a Voronoi diagram and a (positive or negative) MDG field. When a voxel $\mathbf{v}$ is processed, the Voronoi diagram is locally reconstructed by adding $\mathbf{v}$ into the Voronoi sites. Then, we update the MDG field by processing voxels within the Voronoi cell $V(\mathbf{v})$ of $\mathbf{v}$.

The reason for ignoring the region outside $V(\mathbf{v})$ is given as follows. Suppose we process voxels in intensity-decreasing order. Let $\mathbf{x}$ be a voxel in another Voronoi cell $V(\mathbf{v}')$. Then, based on the principle of Voronoi diagram, we have

$$|\mathbf{v} - \mathbf{x}| > |\mathbf{v}' - \mathbf{x}|.$$

On the other hand, since $\mathbf{v}'$ has already been processed and added into the Voronoi sites prior to $\mathbf{v}$, and $\mathbf{x}$ has not been processed yet, we have

$$I(\mathbf{v}') \geqslant I(\mathbf{v}) \geqslant I(\mathbf{x}).$$

Therefore, we have

$$|\nabla_d I(\mathbf{v}', \mathbf{x})| > |\nabla_d I(\mathbf{v}, \mathbf{x})|,$$

and there is no need to update the MDG vector for $\mathbf{x}$ after adding $\mathbf{v}$.

Similarly, in intensity increasing order, since $I(\mathbf{v}') \leqslant I(\mathbf{v}) \leqslant I(\mathbf{x})$, we also have

$$|\nabla_d I(\mathbf{v}', \mathbf{x})| > |\nabla_d I(\mathbf{v}, \mathbf{x})|.$$

Finally, to derive the signed-MDG field, the obtained positive and negative components are combined together according to their absolute values. Specifically, at voxel $\mathbf{v}$, the signed-MDG field is obtained by comparing the absolute values of the magnitude of the positive and negative MDG fields at $\mathbf{v}$, and it is set to the one, which has the larger absolute magnitude value. Note that in practice the image volume is always preprocessed by convolving with a Gaussian kernel before computation to minimize the effect of noise and artifacts. The two Voronoi based processing steps are terminated when all voxels are processed. A flow chart for summarizing the above process is shown in Fig. 10.

4.2. Multi-resolution image registration

4.2.1. Estimation of probability distribution

To compute the multi-dimensional mutual information in our proposed similarity metric (see Eq. (9)), three joint probability distributions need to be estimated (see Section 3.1). For computational efficiency, the estimation is obtained by the normalization of joint histograms rather than Parzen windowing (Bishop, 1995).

Given a transformation registration $\mathbf{T}$, samples are taken from the whole volume or a subset of the floating image f, and then transformed by $\mathbf{T}$ into the reference image r. For the construction of joint histograms in our implementation, the sampling set $\mathcal{V}$ is not restricted within the volume of overlap of f and r. The work by August and Kanade (2005) and also our empirical experience show that
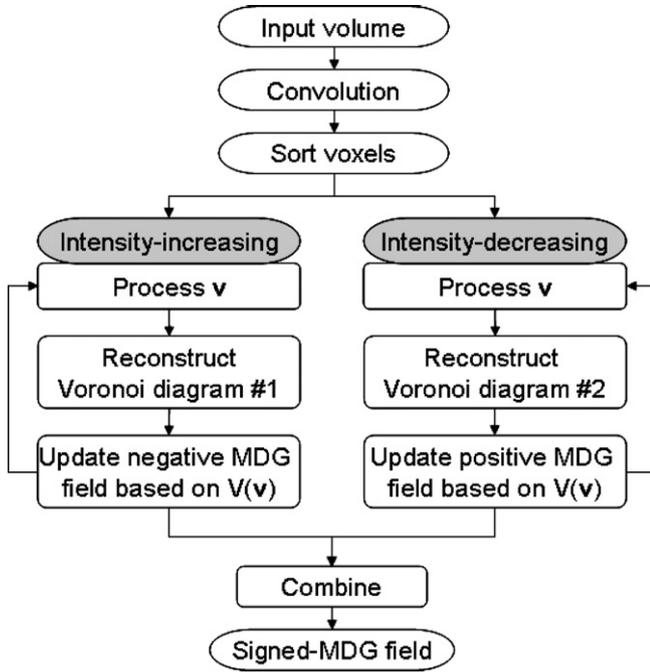
Fig. 10. The flow chart of computing the signed-MDG field.

$$p_{\mathbf{T}}(I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}_r\|) = \frac{h_{\mathbf{T}}(I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}_r\|)}{\sum\limits_{I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}r\|} h_{\mathbf{T}}(I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}_r\|)},$$

$$p_{\mathbf{T}}(I_f, \|\widehat{G}_f\|) = \sum\limits_{I_r, \|\widehat{G}_r\|} p_{\mathbf{T}}(I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}_r\|),$$

$$p_{\mathbf{T}}(I_r, \|\widehat{G}_r\|) = \sum\limits_{I_f, \|\widehat{G}_f\|} p_{\mathbf{T}}(I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}_r\|).$$

4.2.2. Multi-resolution optimization

In order to accelerate the registration process, a multi-resolution approach based on the Gaussian Pyramid representation (Burt and Adelson, 1983; Cole-Rhodes et al., 2003; Wells et al., 1996) is exploited. Rough estimates of $\widehat{\mathbf{T}}$ can be found using downsampled images with Eq. (10), and then treated as starting values for optimization at higher resolutions. Then the fine-tuning of the solution can be derived at the original image resolution.

In our proposed registration method, three resolution levels are used and the definition of resolution levels in the Gaussian Pyramid representation follows the same convention as in Burt and Adelson (1983), i.e., Level 0 image represents the highest and most original resolution and Level 2 image represents the lowest resolution. Smoothing is performed via the binomial filter with coefficients [1,4,6,4,1] (Wells et al., 1996). For easier implementation, all voxels in the floating volume are used to construct the joint histograms at all resolutions (see the above subsection for details). For optimization at each resolution, we use the Powell's direction set method (Press et al., 1992) with Brent's 1D line minimization, where the fractional precision convergence parameters for the Powell and Brent methods are set to 10^{-4} and 10^{-3}, respectively.

5. Experimental results and discussions

To demonstrate the proposed similarity metric defined as Eq. (9) (hereafter referred to as **MI4d*cos**) and the proposed multi-resolution registration method, we performed a set of rigid image registration experiments on X-ray computed-tomography (CT) and T1-weighted MRI (T1). Besides our method, three other similarity measures and registration methods in multi-resolution contexts were also tested for comparison:

the regions from the non-overlapping volume play an important role in information theoretic registration, and by including them can increase the registration robustness. Therefore, for all entropy-based similarity measures we use (e.g. the proposed multi-dimensional mutual information, the conventional mutual information, and other two measures tested in Section 5 for a comparative study), we construct the probability distributions by sampling the whole volume or a subset of f for a fair comparison.

By binning the feature vector pairs, $\mathbf{Z}_f(\mathbf{v}) = (I_f(\mathbf{v}), \|\widehat{G}_f(\mathbf{v})\|)$ and $\mathbf{Z}_r(\mathbf{T}(\mathbf{v})) = (I_r(\mathbf{T}(\mathbf{v})), \|\widehat{G}_r(\mathbf{T}(\mathbf{v}))\|)$, for all $\mathbf{v} \in \mathcal{V}$, a four-dimensional joint histogram, $h_{\mathbf{T}}(I_f, \|\widehat{G}_f\|, I_r, \|\widehat{G}_r\|)$, is then constructed. Below are some implementation details for the construction of histograms in our work:

- The trilinear partial volume distribution interpolation (Maes et al., 1997) is exploited to update the joint histograms for non-grid alignment.
- When $\mathbf{T}(\mathbf{v})$ falls outside the volume of r, a feature vector randomly selected from the background of r is used to update the joint histograms.
- Due to the relatively high dimensionality, the number of four-dimensional joint histogram bins should be limited. In practice, we have found that histograms with 32 bins both for the intensity and magnitude of signed-MDG dimensions can give good performance for registering two image volumes of size $256 \times 256 \times 26$. (Note that the number of histogram bins may be tuned for downsampled images in the multi-resolution registration process.)

We estimate the probability distributions mentioned in Section 3.1 as follows,

- The conventional mutual information measure on image intensity (Maes et al., 1997; Wells et al., 1996) (hereafter referred as **MI2d**).
- Adaptation of mutual information by incorporating spatial information proposed by Pluim et al. (2000) (hereafter referred as **Pluim**), which is the multiplication of the mutual information and a gradient term based on both the magnitude and orientation of image gradient.

- Second-order mutual information proposed by Rueckert et al. (2000) (hereafter referred as **Rueckert**), which considers the co-occurrence of intensity pairs of adjacent points in each of the images to be registered.

The study on the behavior of four registration functions with regard to capture range[2] is presented in Section 5.2. Section 5.3 shows the performance comparisons on registration robustness and accuracy between the proposed method and other three methods.

5.1. Image datasets and ground truth

In the experiments described below, a CT–T1 registration problem was considered. A set of clinical CT–T1 data were obtained from the RIRE project. Note that all the T1 images have been rectified for intensity inhomogeneity and scaling. In general, the size of a CT image volume is $512 \times 512 \times 30$ voxels and the voxel size is $0.65 \times 0.65 \times 4 \text{ mm}^3$, and a T1 image contains $256 \times 256 \times 26$ voxels of dimensions $1.25 \times 1.25 \times 4 \text{ mm}^3$.

With regard to the CT–T1 data, we determined the "ground truth" for registration experiments as follows. First, the multi-resolution mutual information based and normalized mutual information (NMI) (Studholme et al., 1999) based methods were used to register the image pairs without incorporating the spatial information. The evaluations of accuracy, measured as the target registration errors (TREs) in ten volumes of interest (VOIs) (West et al., 1997), were then obtained from the RIRE project. By examining the median errors, four datasets (Datasets pt-001, pt-003, pt-005 and pt-007) with less than 1mm registration error were selected and used in the experiments. Then, the corresponding optimal transformations, whose median errors were 0.5077 mm (for pt-001), 0.5194 mm (for pt-003), 0.7807 mm (for pt-005) and 0.3254 mm (for pt-007) respectively, were used as the ground truth registrations.

5.2. Behavior of CT–T1 registration functions

To study the behavior of the four similarity measures, i.e., **MI2d**, **Pluim**, **Rueckert**, and **MI4d*cos**, we compared their capture ranges along different axes and at different resolutions by using the probing tests.

Fig. 11 plots the probing curves in three translational (the upper half) and three rotational (the lower half) axes (i.e., translation coronal axis, translation sagittal axis, translation transversal axis, rotation coronal axis, rotation sagittal axis, rotation transversal axis) for registering a low resolution (Level 2) CT–T1 image pair (pt-001).

We assume two images are pre-aligned before probing. The first and third columns in the figure demonstrate that obvious local maxima occur for both **MI2d** and **Rueckert**, when the in-plane translational misalignment of two images is relatively large. On the contrary, as observed in the second and fourth columns in the figure, the shape of the probing curves along in-plane translational axes based on **Pluim** and **MI4d*cos** is significantly improved, with noting that those local maxima almost or entirely disappear. Moreover, it is shown that the **MI4d*cos** value keeps decreasing constantly as the increase of the in-plane translational offset. On the other hand, the **Pluim** value is almost constant when the in-plane translational offset is relatively large. This observation suggests that the capture ranges of **MI4d*cos** along in-plane translational axes can be longer than those of the other three measures. This is because:

- With the help of the proposed signed-MDG field, regions with homogeneous intensities (including the anatomical structures and background regions) can provide useful and varying information related to (1) the intensity difference, (2) the distance and (3) the direction of a voxel to a MDG source point. Therefore, when the in-plane misalignment increases, the **MI4d*cos** values would keep decreasing. We believe that such a behavior can benefit the optimization procedure for registration and hence increase the registration robustness.
- For the other three measures, relatively less or even no spatial information has been considered:
  - For **MI2d**, no spatial information is taken into account.
  - For **Rueckert**, only one adjacent neighbor for each voxel is considered at a time.
  - For **Pluim**, although strong gradient information in both images has been incorporated, plenty of spatial information from homogeneous regions is discarded.

On the other hand, for the three rotational probes, the capture ranges of the four similarity measures are comparable (see the lower half of Fig. 11).

As a further study on the behavior of the four objective functions, Fig. 12 plots the probing curves in an in-plane translational axis for registering a CT–T1 image pair (pt-001) at the original resolution (Level 0).

A similar observation that **MI4d*cos** gives a longer capture range in the in-plane translational axis than **MI2d**, **Pluim** and **Rueckert** can be found.

It is worth noting that, although different anatomical details are depicted in CT and T1 image volumes, as a preliminary indicator, these probing tests demonstrate the superior behavior of the proposed **MI4d*cos** measure for CT–T1 registration. We may view this observation as an empirical justification that the proposed method has longer capture range than the other compared methods.

---

[2] Capture range represents the range of alignments from which a registration function can converge to the correct optimum.
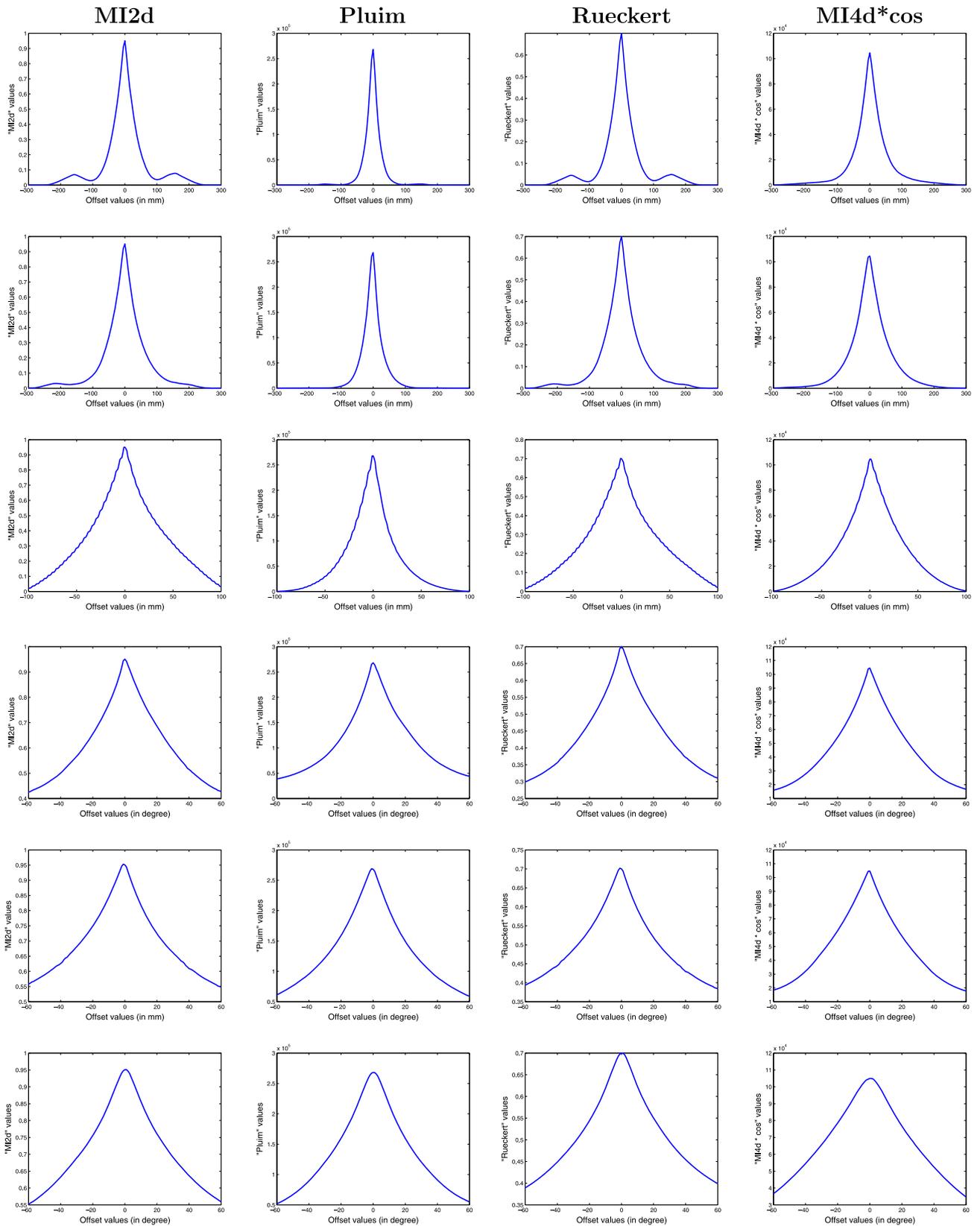
Fig. 11. Probing curves for 3D–3D registration on a CT–T1 image pair (pt-001) at the low resolution (Level 2) in three translational and three rotational axes. From the top row to the bottom row: translation coronal axis, translation sagittal axis, translation transversal axis, rotation coronal axis, rotation sagittal axis, rotation transversal axis.
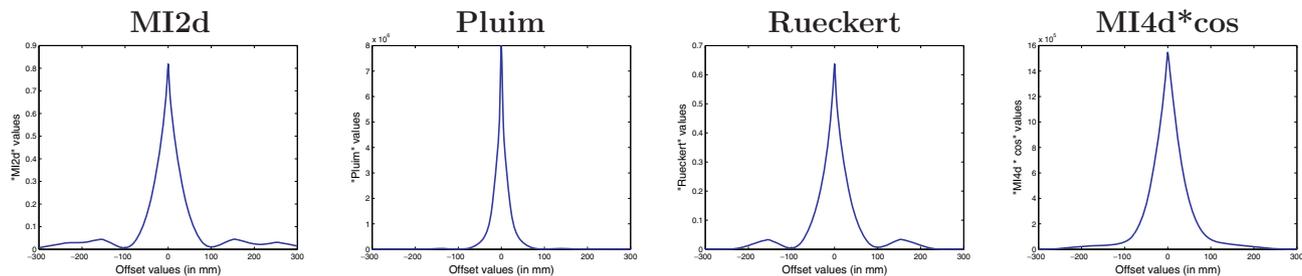
Fig. 12. Probing curves for 3D–3D registration on a CT–T1 image pair (pt-001) at the original resolution (Level 0) along the translation coronal axis.

5.3. Performance comparisons on registration robustness and accuracy

In order to study and compare the registration robustness of four registration methods based on **MI2d**, **Pluim**, **Rueckert** and **MI4d*cos**, a series of randomized experiments has been designed and performed on the aforementioned four CT–T1 image pairs from the RIRE project, i.e., Datasets pt-001, pt-003, pt005 and pt-007 with all resolution levels.

The experiments took 100 tests on each testing image pair for each method. At each trial, the pre-obtained ground truth registration parameters (see Section 5.1) of the testing image pair were perturbed by six uniformly distributed random offsets for all translational and rotational axes. The perturbed parameters were then treated as the starting alignment for registration. In order to show the high optimization capability of **MI4d*cos** with respect to initial alignment, random offsets were generated as follows: (1) For translation coronal axis, offsets were drawn between around $[-125, 125]$ mm; (2) For translation sagittal axis, offsets were drawn between around $[-160, 160]$ mm; (3) For translation transversal axis, offsets were drawn between around $[-70, 70]$ mm; (4) For each rotational axes, offsets were drawn between $[-0.4, 0.4]$ radians, i.e., $[-23, 23]$ degrees. (Note that the ranges for translational offsets were set so that two brains in CT and T1 images have at least 20% overlapping region.) For fair comparison, the same set of randomized starting alignments was used for all **MI2d**, **Pluim**, **Rueckert** and **MI4d*cos**.

To evaluate each obtained registration result with respect to the corresponding ground truth registration, similar to Maes et al. (1997) and Knops et al. (2006), a tight bounding box was fitted around the brain for each T1 image. For each of the eight corner points, the Euclidean distance between the ground truth position and the position transformed by the solution was computed. The median value of the eight distances was then taken for assessing the registration success. A registration was judged to be *successful* if the median error was smaller than or equal to 4 mm, which was the largest voxel dimension of the CT–T1 image pair; otherwise, it was considered a misregistration. For illustration, a pair of successfully registered 2D image slices is shown in Fig. 13.
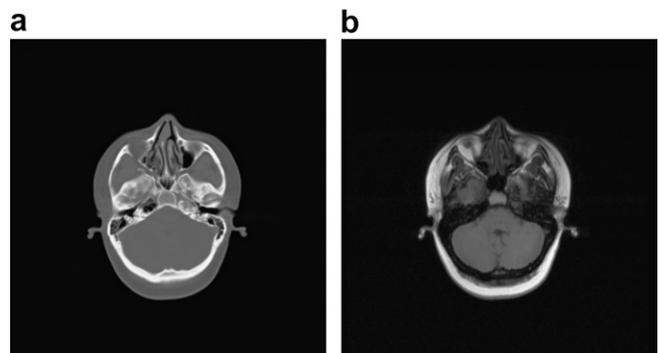

Fig. 13. A pair of successfully registered 2D image slices.

The success rates of the multi-resolution registration methods based on the four similarity measures on the four CT–T1 image pairs are listed in Table 1. It is suggested that: (1) The methods based on **MI2d** and **Rueckert** produce similar success rates; (2) The success rates for **Pluim** are higher than those for **MI2d** by around $10\% \sim 15\%$; (3) The **MI4d*cos** based method consistently gives the highest success rates among all methods for all testing image pairs, specifically, higher than those of the other three methods by about 30%.

Furthermore, based on these experiments, it is noticed that the majority of *unsuccessful* cases for the **MI4d*cos** based method had large misalignments (i.e., about 180°) in one or two rotational axes, while registration errors for other axes were quite small. That is to say, after registration, the brain in the floating image was just inverted along one or two rotational axes. (Hereafter we call such a registration result a *rotational failure*.) More specifically, the numbers of rotational failures for **MI4d*cos** are: 7 for pt-001, 7 for pt-003, 6 for pt-005 and 6 for pt-007, respectively. This observation may imply that the capture ranges for **MI4d*cos** along the translational axes can be relatively long and the proposed method is more robust to translational misalignment.

To compare the registration accuracy of the four registration methods, Table 2 lists the means and standard deviations of the median errors (in millimeters) for the successful registrations for each testing image pair, which are obtained in the experiments discussed above. It is observed that **MI2d** and **Pluim** produce slightly lower med-

Table 1

The success rates of the multi-resolution registration methods based on the five similarity measures on the four CT–T1 image pairs

	Success rate				
	MI2d (%)	**Pluim** (%)	**Rueckert** (%)	**MI4d*cos** (%)	**MI4d** (%)
pt-001	56	69	57	94	56
pt-003	51	64	55	93	67
pt-005	58	67	55	95	71
pt-007	52	64	49	93	70

Table 2

The means and standard deviations of the median errors (in millimeters) for successful registrations for different CT–T1 image pairs

	Registration errors: mean ± sd (in mm)				
	MI2d	**Pluim**	**Rueckert**	**MI4d*cos**	**MI4d**
pt-001	0.6134 ± 0.0525	0.6245 ± 0.1505	0.8532 ± 0.0384	1.3150 ± 0.1401	0.9527 ± 0.1802
pt-003	0.8492 ± 0.0747	0.4829 ± 0.0783	1.1934 ± 0.4096	0.8015 ± 0.2528	0.7931 ± 0.8036
pt-005	0.7353 ± 0.3608	0.5236 ± 0.5720	1.0636 ± 1.0425	1.2993 ± 1.0524	1.1692 ± 0.3597
pt-007	1.2950 ± 1.0485	1.2895 ± 0.8052	1.0952 ± 0.2633	1.2757 ± 0.2780	1.2107 ± 1.0273

ian errors than **Rueckert** and **MI4d*cos**, while, the registration accuracy of the four methods are high in general. Please note that these errors may not be exactly accurate and objective, since we lack the gold standard registrations, which remain sequestered in the RIRE project. The determination of the ground truth registrations, which we used as reference, can be found in Section 5.1.

To study the impact of including orientation information in the proposed similarity measure, the orientation component in our registration objective function is removed, i.e. $S(\mathbf{T}) = MI(\mathbf{T})$ (hereafter referred as **MI4d**) and we have performed another set of randomized rigid registration experiments for each pair of images. The results are listed in the last columns of Tables 1 and 2. By comparing **MI4d*cos** and **MI4d**, it is observed that the robustness can be improved by incorporating the orientation information of the MDG vector field in our proposed metric, and the accuracies are comparable.

The average running times for **MI2d**, **Pluim**, **Rueckert**, **MI4d**, and **MI4d*cos** were 42 s, 4 min 36 s, 12 min 18 s, 11 min 51 s, and 28 min 3 s, respectively. All the experiments were conducted in a PC with 3.19 GHz CPU and 1GB RAM.

6. Conclusion

To overcome the lack of concern on the spatial information in the conventional mutual information based image registration framework, this paper has designed a new spatial feature field, namely the maximum distance-gradient (MDG) vector field, for registration tasks. The MDG vector field encodes both local edge information and globally defined spatial information, which includes (1) the intensity difference, (2) the distance and (3) the direction of a voxel to a MDG source point. By considering the intensity difference between a voxel and its MDG source, we have further extended it to become the signed-MDG feature field.

In order to include the magnitude and orientation information of the signed-MDG feature field into the conventional mutual information based registration framework, our overall similarity metric is the combination (i.e., multiplication) of two separate measure components. The first component is the multi-dimensional mutual information measure on a two-element attribute vector space, which is formed by associating the magnitude of the signed-MDG feature with the image intensity. The second component is based on an angle measure on the orientation of the MDG field of the two images to be registered. To increase computational efficiency and robustness of the proposed method, the overall registration procedure has been a multi-resolution iterative process.

The experimental results on clinical three-dimensional CT and MR-T1 datasets from the RIRE project have indicated that, at different image resolutions, the proposed registration function has longer capture ranges than the other three related similarity measures in the comparison. Moreover, a large number of (around 2000) randomized experiments on precisely registered clinical CT–T1 image pairs have demonstrated that the success rates of our method are consistently and significantly higher than those of the other three methods. It has also been shown that the registration accuracy of the new method is high.

References

August, J., Kanade, T., 2005. The role of non-overlap in image registration. Information Processing in Medical Imaging, 19th International Conference. IPMI, pp. 713–724.

Bardera, A., Feixas, M., Boada, I., Sbert, M., 2006. High-dimensional normalized mutual information for image registration using random lines. In: Pluim, J.P.W., Likar, B., Gerritsen, F.A. (Eds.), WBIR. Lecture Notes in Computer Science, vol. 4057. Springer, pp. 264–271.

Bishop, C., 1995. Neural Networks for Pattern Recognition. Oxford University Press.

Burt, P., Adelson, E., 1983. The Laplacian pyramid as a compact image code. IEEE Transactions on Communications 31 (4), 532–540.

Butz, T., Thiran, J.-P., 2001. Affine registration with feature space mutual information. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 549–556.

Cole-Rhodes, A., Johnson, K., LeMoigne, J., Zavorin, I., 2003. Multi-resolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. IEEE Trans. Image Process. 12 (12), 1495–1511.

Cover, T., Thomas, J., 1991. Elements of Information Theory. John Wiley & Sons, Inc., New York.

Knops, Z., Maintz, J., Viergever, M., Pluim, J., 2006. Normalized mutual information based registration using k-means clustering and shading correction. Medical Image Analysis 10 (3), 432–439.

Ma, B., Hero, A.O., Gorman, J., Michel, O., 2000. Image registration with minimum spanning tree algorithm. In: International Conference on Image Processing, vol. 1.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imaging 16 (2), 187–198.

Meyer, C.R., Boes, J.L., Kim, B., Bland, P.H., Zasadny, K.R., Kison, P.V., Koral, K., Frey, K.A., Wahl, R.L., 1997. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. Med. Image Anal. 1 (2), 195–206.

Neemuchwala, H., Hero, A., Carson, P., 2005. Image matching using alpha-entropy measures and entropic graphs. Signal Process. 85 (2), 277–296.

Penney, G., Weese, J., Little, J., Desmedt, P., Hill, D., Hawkes, D., 1998. A comparison of similarity measures for use in 2D–3D medical image registration. IEEE Trans. Med. Imaging 17 (4), 586–595.

Pluim, J., Maintz, J., Viergever, M., 2000. Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imaging 19 (8), 809–814.

Pluim, J., Maintz, J., Viergever, M., 2003. Mutual-information-based registration of medical images: a survey. IEEE Trans. Med. Imaging 22 (8), 986–1004.

Preparata, F.R., Shamos, M.I., 1985. Computational Geometry: An Introduction. Springer-Verlag, New York.

Press, W., Teukolsky, S., Vetterling, W., Flannery, B., 1992. Numerical Recipes in C, second ed. Cambridge University Press.

Rueckert, D., Clarkson, M., et al., 2000. Non-rigid registration using higher-order mutual information. In: SPIE, Medical Imaging: Image Processing, vol. 3979, pp. 438–447.

Russakoff, D.B., Tomasi, C., Rohlfing, T., Maurer, Jr., C.R., 2004. Image similarity using mutual information of regions. In: Pajdla, T., Matas, J. (Eds.), ECCV (3). Lecture Notes in Computer Science, vol. 3023. Springer, pp. 596–607.

Sabuncu, M.R., Ramadge, P.J., 2003. Spatial information in entropy-based image registration. In: Gee, J.C., Maintz, J.B.A., Vannier, M.W. (Eds.), WBIR. Lecture Notes in Computer Science, vol. 2717. Springer, pp. 132–141.

Spivak, M., 1980. Calculus, second ed. Publish or Perish, Berkeley, CA.

Studholme, C., Hill, D., Hawkes, D., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recog. 32, 71–86.

Studholme, C., Hill, D.L.G., Hawkes, D.J., 1997. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. Med. Phys. 24 (1), 25–35.

Vajda, I., 1989. Theory of Statistical Inference and Information. Kluwer Academic Publishers, Dordrecht.

Wells, W., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R., 1996. Multi-modal volume registration by maximization of mutual information. Med. Image Anal. 1 (1), 35–51.

West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer, Jr., C.R., Kessler, R.M., Maciunas, R.J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., van den Elsen, P.A., Napel, S., Sumanaweera, T.S., Harkness, B., Hemler, P.F., Hill, D.L.G., Hawkes, D.J., Studholme, C., Maintz, J.B.A., Viergever, M.A., Malandain, G., Pennec, X., Noz, M.E., Maguire, Jr., G.Q., Pollack, M., Pelizzari, C.A., Robb, R.A., Hanson, D., Woods, R.P., 1997. Comparison and evaluation of retrospective intermodality brain image registration techniques. J. Comput. Assist. Tomogr. 21 (4), 554–566.