CrossMark

ORIGINAL ARTICLE

Multimodal image registration with joint structure tensor and local entropy

Jingya Zhang[1,2] · **Jiajun Wang**[1] · **Xiuying Wang**[3] · **Dagan Feng**[3]

Abstract

Purpose Nonrigid registration of multimodal medical images remains a challenge in image-guided interventions. A common approach is to use mutual information (MI), which is robust to the intensity variations across modalities. However, primarily based on intensity distribution, MI does not take into account of underlying spatial and structural information of the images, which might lead to local optimization. To address such a challenge, this paper proposes a two-stage multimodal nonrigid registration scheme with joint structural information and local entropy.

Methods In our two-stage multimodal nonrigid registration scheme, both the reference image and floating image are firstly converted to a common space. A unified representation in the common space for the images is constructed by fusing the structure tensor (ST) trace with the local entropy (LE). Through the representation that reflects its geometry uniformly across modalities, the complicated deformation field is estimated using L_1 or L_2 distance.

Results We compared our approach to four other methods: (1) the method using LE, (2) the method using ST, (3) the method using spatially weighted LE and (4) the conventional MI-based method. Quantitative evaluations on 80 multimodal image pairs of different organs including 50 pairs of MR images with artificial deformations, 20 pairs of medical brain MR images and 10 pairs of breast images showed that our proposed method outperformed the comparison methods. Student's *t* test demonstrated that our method achieved statistically significant improvement on registration accuracy.

Conclusion The two-stage registration with joint ST and LE outperformed the conventional MI-based method for multimodal images. Both the ST and the LE contributed to the improved registration accuracy.

✉ Jiajun Wang
  jjwang@suda.edu.cn

  Jingya Zhang
  zhangjy0611@163.com

  Xiuying Wang
  xiu.wang@sydney.edu.au

  Dagan Feng
  dagan.feng@sydney.edu.au

[1] School of Electronic and Information Engineering, Soochow University, Suzhou 215006, People's Republic of China

[2] Department of Physics, Changshu Institute of Technology, Changshu 215500, People's Republic of China

[3] Biomedical and Multimedia Information Technology Research Group, School of Information Technologies, University of Sydney, Sydney, NSW 2006, Australia

Introduction

Biomedical images are indispensable to improve the effectiveness and safety of modern radiology and surgery. Medical images often from multiple imaging modalities, collected at various times, are processed and analyzed for more accurate diagnosis and treatment planning of a wide range of cancers [1,2]. Medical image registration that creates alignment and correspondence between the anatomical or functional regions of different images is essential in adaptive image-guided radiation therapy. Registration of the planning images with the onsite images is critical to adjust radiation treatment and dose delivery according to the changes introduced by patient movements, the tumor regression or progres-

sion, and invulnerable movement of the surround organs [3,4].

Generally, image registration can be broadly divided into feature-based and intensity-based techniques. Feature-based methods rely on extracting and matching features of the input image pair, such as contours, corners or manual markers. Such methods are challenged in matching images from different modals although they are often more efficient to compute. Intensity-based registrations directly measure the degree of shared intensity information between images [5,6]. This category of methods optimizes the similarity measure between the floating and reference image [7–9] via searching the optimal transformation. Mutual information (MI) is the most recognized metric for multimodal registration due to its ability to handle large intensity variations [10]. However, it is worth noting that MI-based registration only takes intensity distributions into account and does not consider the underlying structure information, which might lead to local optimization.

Structural representation for multimodal images has gained great interest for multimodal registration. For example, local patch-based entropy images are structural representations for multimodal images [11]. However, there might be the possible ambiguities when several patches share a same entropy value. Wachinger et al. [11] computed a location-dependent weighting to address this issue. Structure tensor analysis is considered to be a way of describing the image structure, which has been applied to inspect and quantify the tissue microstructure presented in diffusion tensor imaging (DTI) images [12]. Although the structure tensor trace can obtain the structural representation of medical images and employs gray values instead of only binary values, however, much internal information might not be preserved.

To address the above-mentioned problems, we propose to construct a structural descriptor which fuses the local entropy with a novel structure tensor trace using an integral image-based filter to describe the geometric and structural properties of data. With the unified structural representation, the consequent nonrigid registration using a simple $L2$ distance similarity is performed on the continuum mechanics constraint.

Methods and materials

As illustrated in Fig. 1, our proposed nonrigid registration method includes two major components, namely structural representation and estimation of geometric mapping. In structural representation process, the reference image and floating image are converted to the unified representation using the trace of image structure tensor and the local entropy. In the second process, the displacement field for the converted image pair generated is estimated based on the continuum mechanics model.

Structural representation based on structure tensor

In order to unify image modalities for registration, the multimodal images are firstly converted to structural representations. Generally, direct use of image gradient would not be suitable across different modalities. However, the use of the local gradient information upon neighborhood [13] may provide a better solution. As a robust gradient method, the structure tensor analysis is an good option for tracking image microstructure; however, this method is problematic in low-resolution images and is sensitive to image noises. To tackle such a problem, we propose to combine the structure tensor trace with the integral image average filter for such a conversion. Such a strategy treats regions, edges, corners and textures in a unified manner and is thus more meaningful than only using intensities.

Average filter on integral images

Integral image can be computed in a recursive way to improve the calculation efficiency [14]. The entry of the integral image $I_\Sigma(\mathbf{x})$ at point $\mathbf{x} = (x, y)$ represents the sum of all pixels in a rectangular region formed by the origin and $\mathbf{x}$ in the input image I.

$$I_\Sigma(\mathbf{x}) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \tag{1}$$

The average filtered result on the integral images can be computed using other three additions (see Fig. 2) and four memory accesses to calculate the mean value of intensity inside a rectangular region of any size. Once the integral image has been computed, the sum of the intensities inside the rectangular region Ω can be calculated as follows:

$$S_\Omega = I_\Sigma(\mathbf{x}_a) - I_\Sigma(\mathbf{x}_b) - I_\Sigma(\mathbf{x}_c) + I_\Sigma(\mathbf{x}_d) \tag{2}$$

The average filtered result at point $\mathbf{x}$ can be obtained as follows:

$$I_f(\mathbf{x}) = \frac{S_\Omega}{M \times N} \tag{3}$$

where M, N is the size of the rectangle in x and y direction.

Structure tensor-based descriptor

Mathematically, the structure tensor $\mathbf{T_k}$ of an image I_f at position $\mathbf{x} = (x, y)$ can be defined as:

$$\mathbf{T}_k = \begin{bmatrix} <I_{fx}, I_{fx}>_w & <I_{fx}, I_{fy}>_w \\ <I_{fx}, I_{fy}>_w & <I_{fy}, I_{fy}>_w \end{bmatrix} \tag{4}$$
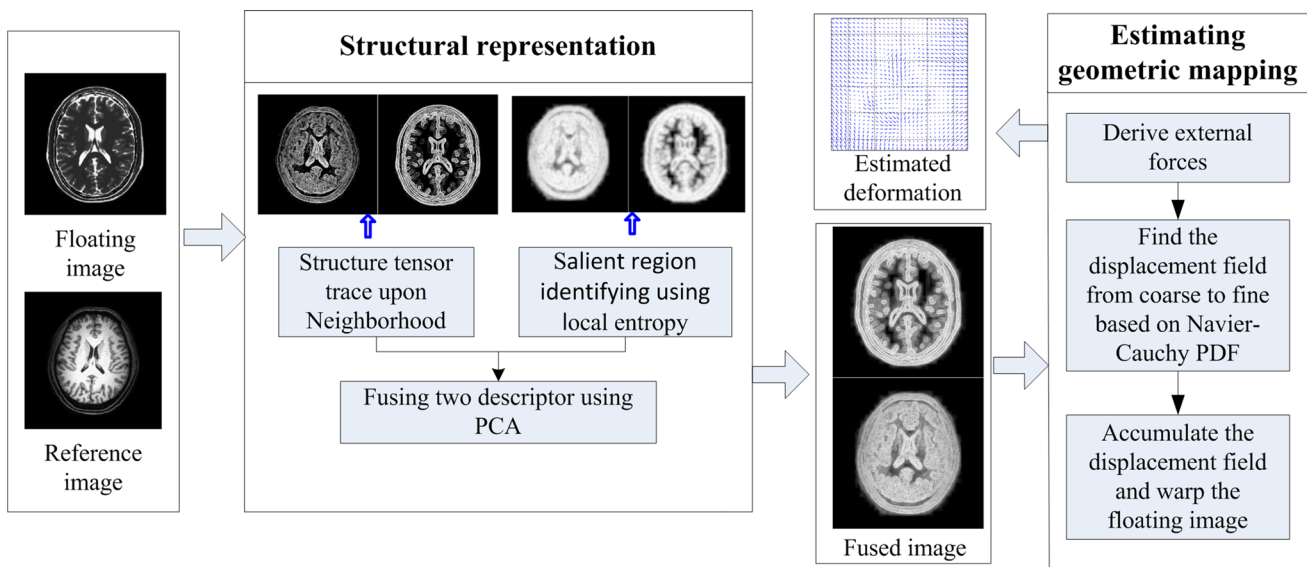
Fig. 1 The framework of our proposed nonrigid registration method
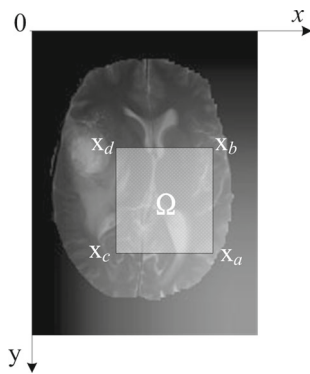

Fig. 2 Illustration of any rectangular Ω in the integral image

where I_{fx} and I_{fy} are, respectively, the partial derivatives of the image I_f computed at position $\mathbf{x} = (x, y)$, $< \cdot, \cdot >_w$ is a weighted inner product operator, e.g.,

$$< I_{fx}, I_{fy} >_w = \iint_{\Re^2} w(x, y) \cdot I_{fx}(x, y) \cdot I_{fy}(x, y) \mathrm{d}x \mathrm{d}y \tag{5}$$

where $w(x, y)$ is a Gaussian function with a specified neighborhood size ($\sigma = 5$ in our paper). The trace value $\mathrm{tr}(\mathbf{T_k})$ of the positive definite second-order tensor $\mathbf{T_k}$ in Eq. (4) represents the structure of the considered image. The larger the value of $\mathrm{tr}(\mathbf{T_k})$ is, the more probable the pixel will be in the image border or corner.

Salient representation based on local entropy

In order to alleviate the information loss problem arising from the structure tensor descriptor, we propose to use the local entropy to compensate the loss of information. Entropy is an important concept for image registration. The widely used mutual information similarity measure for multimodal registration is computed based on the entropy of the joint and marginal probability distribution. The Shannon entropy $H_l(\mathbf{x})$ defined on the neighborhood N_x around $\mathbf{x}$ is calculated as:

$$H_l(X) = -\sum_{i \in I} p(X = i) \cdot \log p(X = i) \tag{6}$$

where X is a discrete random variable (with possible values in I) representing the intensity of pixels in the neighborhood, p is the probability density function (PDF) of X.

Data fusion

We combine the low-resolution local entropy with the higher-resolution structure tensor trace by fusion technique to improve the interpretability of the fused data. Various image fusion techniques are available in published literature. In this study, data fusion is performed on a pixel basis using principal component analysis (PCA) which has been successfully used earlier for fusion of optical and synthetic aperture radar data [15].

Let D be the fusion output which is a linear combination of the structure tensor trace value X_1 and the local entropy X_2. Thus:

$$D = m_1 X_1 + m_2 X_2 \tag{7}$$

where m_1 and m_2 are the associated eigenvector components with the largest eigenvalue of covariance matrix A. Figure 3
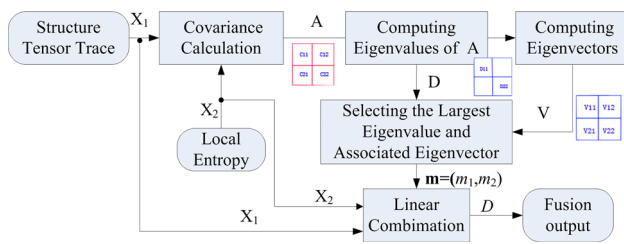
Fig. 3 The process of data fusion with PCA

gives details of the fusion procedure using PCA. The fused image reflects the multifaceted information of the source images and is more suitable for consequent registration.

Multimodal image registration for the converted images

Our registration of two structural representations is achieved by obtaining the displacement field from an elastic model that is based on continuum mechanics. The deformation can be described with the Navier–Cauchy partial differential equation (PDE) [16]:

$$\frac{E}{2(1+\nu)}\nabla^2\mathbf{u} + \frac{E}{2(1+\nu)(1-2\nu)}\nabla(\nabla\cdot\mathbf{u}) + \mathbf{f}(\mathbf{x},\mathbf{u}) = 0$$

$$(8)$$

where E is the Young's modulus, and ν is the Poisson's ratio. $\mathbf{f}(\mathbf{x},\mathbf{u})$ is the external constraint imposed by the image pairs' similarity metric.

Since the intensities of different modalities can be unified as a modality independent descriptor, a variant of the L_2 distance metric, which is often used in mono-modality registration, is employed to derive the external force:

$$\mathbf{f}(\mathbf{x}, \mathbf{u}) = (D_m(\mathbf{x} + \mathbf{u}) - D_r(\mathbf{x})) \cdot \nabla D_m(\mathbf{x} + \mathbf{u}) \qquad (9)$$

where $D_m(\mathbf{x}+\mathbf{u})$ and $D_r(\mathbf{x})$ are, respectively, the descriptors values at points $\mathbf{x} + \mathbf{u}$ and $\mathbf{x}$ of the floating and reference images. $\nabla D_m(\mathbf{x} + \mathbf{u})$ is the gradient of $D_m(\mathbf{x} + \mathbf{u})$.

Materials

To evaluate our registration method based on the fused structural representation (FSR), a realistically generated synthetic brain dataset with registered T1, T2, FLAIR and post-gadolinium T1c MR images from one subject was used to generate image pairs for different methods. The skull in these provided brain MR scans has been wiped off. The dataset was obtained from the brain tumor segmentation (BRATS) challenge [17]. Thirty B-spline-based synthetic deformation fields were generated to deform the T2, FLAIR and post-gadolinium T1c MR images to generate floating images. Meanwhile, the original T1, T2 and FLAIR images were

used as the reference images. We performed nonrigid registration on 50 multimodal image pairs chosen from these reference images and floating images.

We also evaluated the performance of our proposed algorithm on clinical images from different organs including 20 brain MRI image pairs and 10 breast image pairs. The MR brain images from healthy volunteers were used. This database was collected and made available by the CASI-Lab at The University of North Carolina at Chapel Hill and was distributed by the MIDAS Data Server at Kitware, Inc [18]. Images were acquired on a 3T unit under standardized protocols. Images include T1 and T2 acquired at $1 \times 1 \times 1$ mm^3, magnetic resonance angiography (MRA) acquired at $0.5 \times 0.5 \times 0.8$ mm^3 and DTI using 6 directions and a voxel size of $2 \times 2 \times 2$ mm^3. We chose one T1-Flash image as the reference and randomly selected 20 T2 images as floating images. The breast images were acquired from 10 patients who had been treated in the First Affiliated Hospital of Soochow University. The cranio-caudal (CC) view of mammography and the MRI image of every patient were used as the reference image and floating image for registration evaluation.

Ethics statement

The study was carried out according to the Helsinki Declaration and approved by the ethical committee of The University of North Carolina at Chapel Hill. The need for informed consent was waived, because the data used in this study had already been collected for clinical purposes. Furthermore, the present study did not interfere with the treatment of patients, and the database was organized in a way that makes the identification of an individual patient impossible.

Methods for comparison

To evaluate the performance of our method and investigate the contribution of structure tensor trace and local entropy information in the deformable multimodal registration, we compared our registration results with the following four methods: (1) the method using local entropy (LE method); (2) the method using structure tensor (ST method); (3) the method using spatially weighted local entropy (WLE method) proposed in the reference [11]; and (4) the conventional MI-based method (MI method).

Accuracy was measured quantitatively and qualitatively. Quantitative similarity measurements included NC (normalized correlation) and NMI (normalized mutual information), and the mean distance between the ground seeds and corresponding ones after registration. Qualitative assessments include visual inspection on the subtraction images and the checkerboard fusion images between the reference image and the floating image after registration. The fusion checkerboard
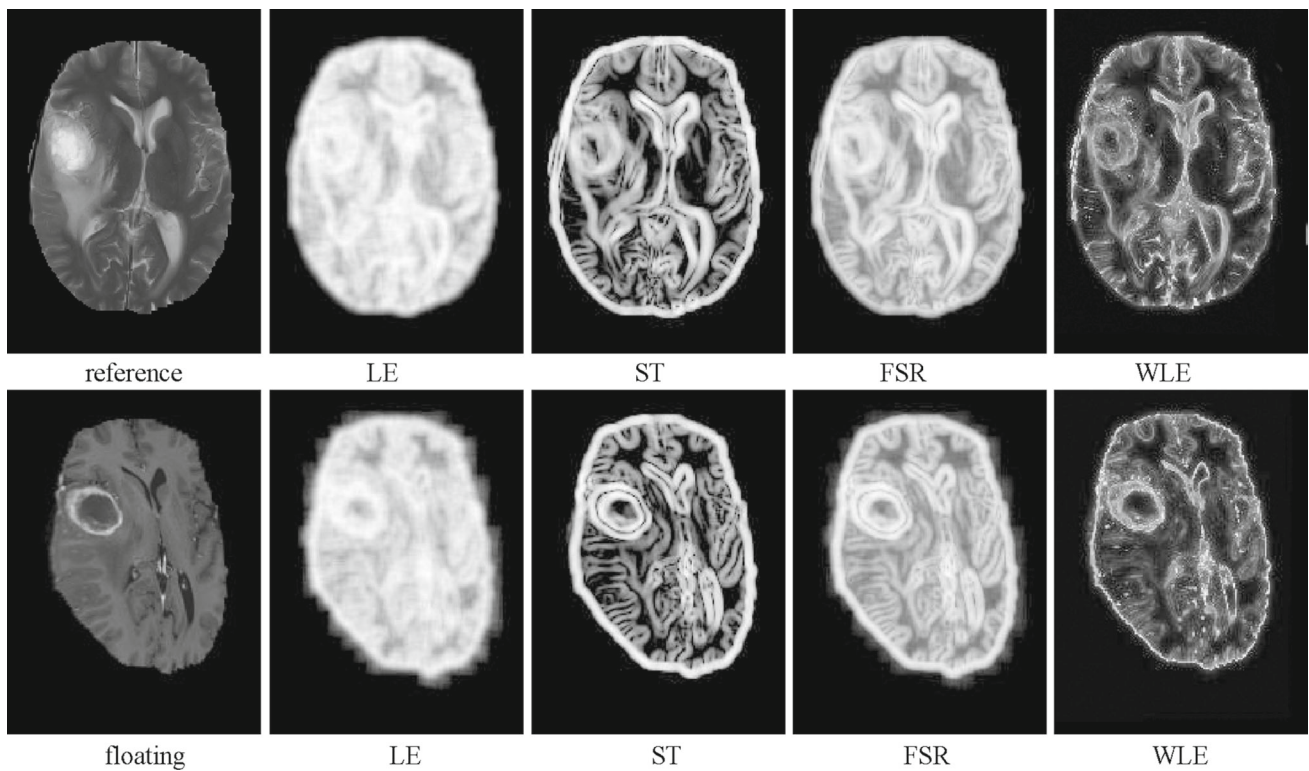
Fig. 4 Examples of MR brain images with artificial deformations. The *top row* shows the reference image of MR-T2 modality and its corresponding representations by local entropy (LE), the structure tensor (ST), the fused structural representation (FSR) and spatially weighted local entropy (WLE). The *bottom row* shows the corresponding floating image of MR-T1c modality and its corresponding structural representations

image highlights the edge continuity of the registration result. The more the continuity in the fusion image, the better the correspondence is achieved from registration.

The checkerboard fusion image is defined as below:

$$I_q = \begin{cases} I_{ref}(x, y), & I_c(x, y) = 255 \\ I_{reg}(x, y), & I_c(x, y) = 0 \end{cases} \quad (10)$$

where $I_q(x, y)$ is the gray value of the fusion image, $I_{reg}(x, y)$ is the gray value of the image after registration, $I_{ref}(x, y)$ is the gray value of the reference image and $I_c(x, y)$ is the gray value of the checkerboard image.

Parameter setting

There are two types of parameters in our method. The first one is the elastic material-related parameters including the elastic modulus E and Poisson's ratio v. As those in [19], we set our model as an isotropic linear elastic model with the elastic modulus $E = 100$kPa and the Poisson's ratio $v = 0.45$. The second one is the parameters for multiresolution registration including the pyramid level number $L(L = 2)$ of levels and the resolution of each level **res**(**res** = {8, 4}).
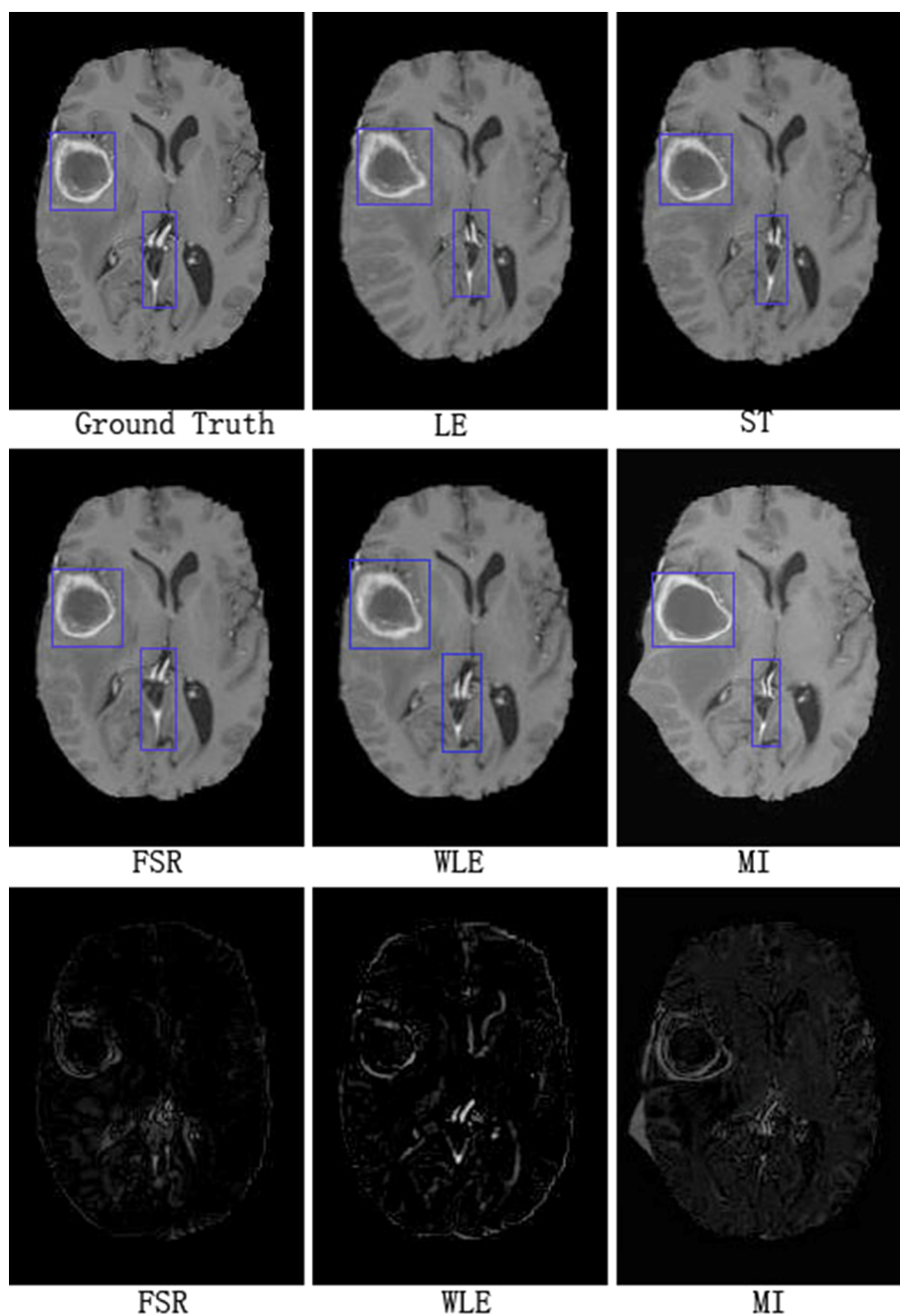
Results

Registration results of MR brain images with artificial deformations

Figure 4 shows the unified structural representations (by LE, ST, FSR and WLE, respectively) of the reference MR-T2 image and the floating MR-T1c image. The floating MR-T1c image was generated by elastically deforming the original MR-T1c image, which was used as a ground truth for evaluating the registration accuracy. With these unified representations, the deformation could be estimated in the common space.

Figure 5 illustrates the registration results of the reference and floating images in Fig. 4 and also the result using MI-based algorithm for visual inspection and evaluation. The image shown in the first column of the top row is the original MR-T1c image used as the ground truth. The subtraction images between the results (by FSR, WLE and MI respectively) and the ground truth are also illustrated in the bottom row of Fig. 5.

Student's *t* tests of the mean distance between the ground seeds and corresponding ones after registration were performed to assess whether the improvement of our FSR

Fig. 5 Registration results of MR brain images with artificial deformations: the *left* column of the *top row* is the original T1c MR image from the dataset used as the ground truth. The registration results, respectively, using LE method, ST method, FSR method, WLE method and MI-based method are shown in the *top row* and *middle row*. Images in the *bottom row* show the subtraction images between the ground truth and result images of FSR, WLE, MI algorithms

method is of statistical importance. The result illustrated in Table 1 shows that at the 0.05 level, the mean values for mean distance of all 50 registration results using FSR method was significantly less than the ones using LE, WLE and MI-based method, while they were marginally less than the ones using the ST method.

Additionally, the comparison results of the other four methods and our FSR algorithm in terms of NC and NMI similarity measurement are given in Tables 2 and 3. The results indicated that at the 0.05 level, the mean values for NC and NMI using our FSR algorithm were significantly greater than

the ones using LE method, WLE and MI method, while they were marginally larger than the ones using the ST method.

Registration results of real inter-subject MR brain images

For the further evaluation of our proposed nonrigid registration algorithm, twenty real MR-T1-Flash and MR-T2 brain images from different subjects were used as the reference and the floating images, respectively. Figure 6 illustrates one example of the inter-subject registration. Figure 6c illus-

Table 1 Mean registration distances (in mm) from different methods [mean value (95 % CI)]

Methods	T1-FLAIR	T1-T1c	T2-FLAIR	T2-T1c	FLAIR-T1c
MI method	0.7852	0.694	0.8806	0.799	0.8024
	(0.6489, 0.9215)	(0.6039, 0.7841)	(0.7918, 0.9693)	(0.5653, 0.8328)	(0.6741, 0.9306)
LE method	0.9361	0.9442	0.9614	0.948	0.8875
	(0.8348, 1.0374)	(0.8021, 1.0862)	(0.8276, 1.0952)	(0.8359, 1.0601)	(0.7253, 1.0497)
ST method	0.6568	0.6043	0.6532	0.6187	0.6667
	(0.5181, 0.7954)	(0.5073, 0.7013)	(0.4999, 0.8065)	(0.5408, 0.6966)	(0.5407, 0.7926)
WLE method	0.8367	0.891	0.8805	0.9419	0.8314
	(0.6483, 1.0251)	(0.7702, 1.0117)	(0.7657, 0.9954)	(0.8865, 0.9973)	(0.6944, 0.9684)
FSR method	0.6369	0.6032	0.6479	0.6111	0.6323
	(0.4976, 0.7761)	(0.4991, 0.7073)	(0.4920, 0.8038)	(0.5253, 0.6970)	(0.5389, 0.7257)
p value (except ST)	<0.05	<0.05	<0.05	<0.05	<0.05

Table 2 Average NC values of different algorithms [mean value (95 % CI)]

Methods	T1-FLAIR	T1-T1c	T2-FLAIR	T2-T1c	FLAIR-T1c
MI method	0.9583	0.9455	0.9743	0.9565	0.9459
	(0.9173, 0.9993)	(0.9421, 0.9489)	(0.9636, 0.9849)	(0.9427, 0.9702)	(0.9322, 0.9596)
LE method	0.9678	0.9551	0.9599	0.9536	0.9472
	(0.9453, 0.9904)	(0.9333, 0.9769)	(0.9434, 0.9765)	(0.9362, 0.9709)	(0.9340, 0.9605)
ST method	0.9784	0.9713	0.9821	0.9774	0.962
	(0.9468, 1.0100)	(0.9506, 0.9921)	(0.9672, 0.9969)	(0.9658, 0.9889)	(0.9487, 0.9753)
WLE method	0.975	0.9571	0.9729	0.9603	0.9567
	(0.9548, 0.9953)	(0.9436, 0.9706)	(0.9609, 0.9849)	(0.9535, 0.9672)	(0.9466, 0.9658)
FSR method	0.9844	0.9736	0.9824	0.9778	0.9646
	(0.9627, 1.0060)	(0.9500, 0.9972)	(0.9677, 0.9972)	(0.9636, 0.9919)	(0.9512, 0.9781)
p value (except ST)	<0.05	<0.05	<0.05	<0.05	<0.05

Table 3 Average NMI values of different algorithms [mean value (95 % CI)]

Methods	T1-FLAIR	T1-T1c	T2-FLAIR	T2-T1c	FLAIR-T1c
MI method	1.3376	1.3235	1.3576	1.3312	1.3301
	(1.2992, 1.3760)	(1.3148, 1.3322)	(1.3438, 1.3714)	(1.3161, 1.3463)	(1.3179, 1.3422)
LE method	1.3569	1.3332	1.3521	1.3315	1.3245
	(1.3333, 1.3805)	(1.3047, 1.3617)	(1.3328, 1.3714)	(1.3079, 1.3551)	(1.3031, 1.3460)
ST method	1.3738	1.3565	1.3814	1.3528	1.3328
	(1.3390, 1.4087)	(1.3239, 1.3890)	(1.3492, 1.4136)	(1.3285, 1.3771)	(1.3040, 1.3616)
WLE method	1.3519	1.3376	1.3524	1.3365	3.3284
	(1.3337, 1.3821)	(1.3130, 1.3623)	(1.3237, 1.3811)	(1.3235, 1.3496)	(1.3134, 1.3434)
FSR method	1.3871	1.3613	1.3887	1.3586	1.3469
	(1.3544, 1.4197)	(1.3251, 1.3975)	(1.3506, 1.4268)	(1.3292, 1.3881)	(1.3257, 1.3681)
p value (except ST)	<0.05	<0.05	<0.05	<0.05	<0.05

trates the estimated deformation field. Figure 6d, e is the registration results from the registration methods based on the MI and our FSR method. Since no ground truth image could be used in the medical inter-subject registration, our visual assessment was performed by observing the enlarged checkerboard fusion of the reference image and the registration results (Fig. 7a, b). In order to assess the continuities of fused images, object contours were also sketched in Fig. 7a,
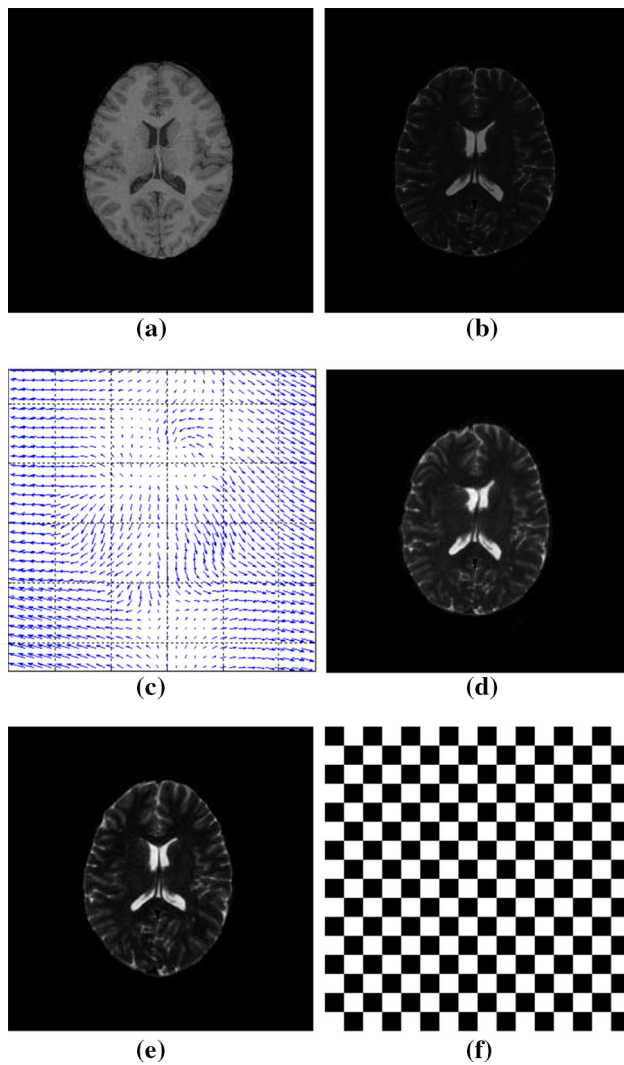
Fig. 6 Registration results of MR-T1 and MR-T2 brain images: **a** reference image, **b** floating image, **c** the estimated deformation field, **d** registrated floating image by MI method, **e** registered floating image by FSR method, **f** checkboard mask

Fig. 7 The checkerboard fusion of the reference with **a** the MI registration result and **b** the FSR registration result. The *circle* region in **b** shows more continuity

Fig. 8 Statistical *box-plots* of the registration results in terms of NMI

b. These images show that the registration result from our FSR method was better aligned to the reference image when compared to the results from MI-based method.

Quantitative accuracy evaluations of twenty image pairs in terms of NMI are given in Fig. 8. As it is shown, the proposed method achieved higher NMI values and in turn further demonstrated that our method outperformed the MI-based method.

Registration results for clinical mammography and MRI images

As a clinical diagnostic tool, mammography is the most commonly used breast imaging modality. Although the resolution of a typical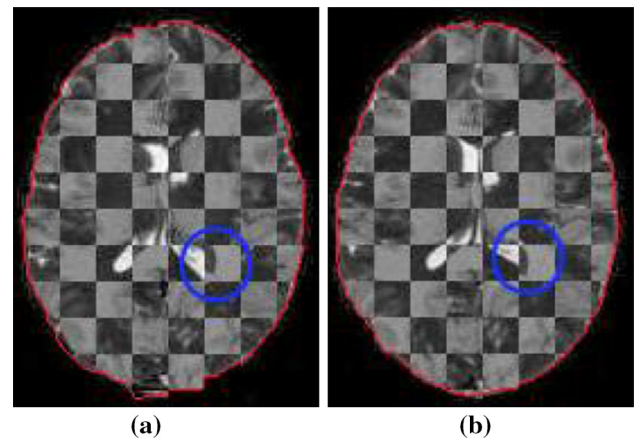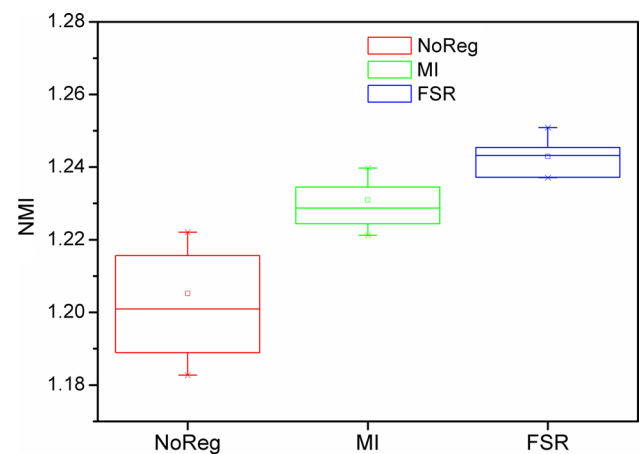 mammograph is reasonable, superimposition of breast tissue and the low contrast between the healthy fibro-glandular and the suspicious lesion make interpretation difficult. MRI is generally used as a complementary modality to compensate the ambiguities in mammography. To fully exploit the complementary information in MRI and mammography, the large and complicated deformation between them should be estimated by the multimodal registration algorithm.

Figure 9 illustrates an example pair of the images used in our registration method evaluation. Figure 9a is a CC mammographic image used as the reference image. Figure 9b is a MRI image used as the floating image acquired from the same patient. Figure 9c is the registration result using MI method, and Fig. 9d is the result using our FSR method. Our visual assessment was performed by observing the overlay of the reference image and the registration result from the floating image. Figure 9e shows the overlay for (a) and (c), and the contour lines with blue and red described the local mismatch between the image pair using MI-based algorithm.
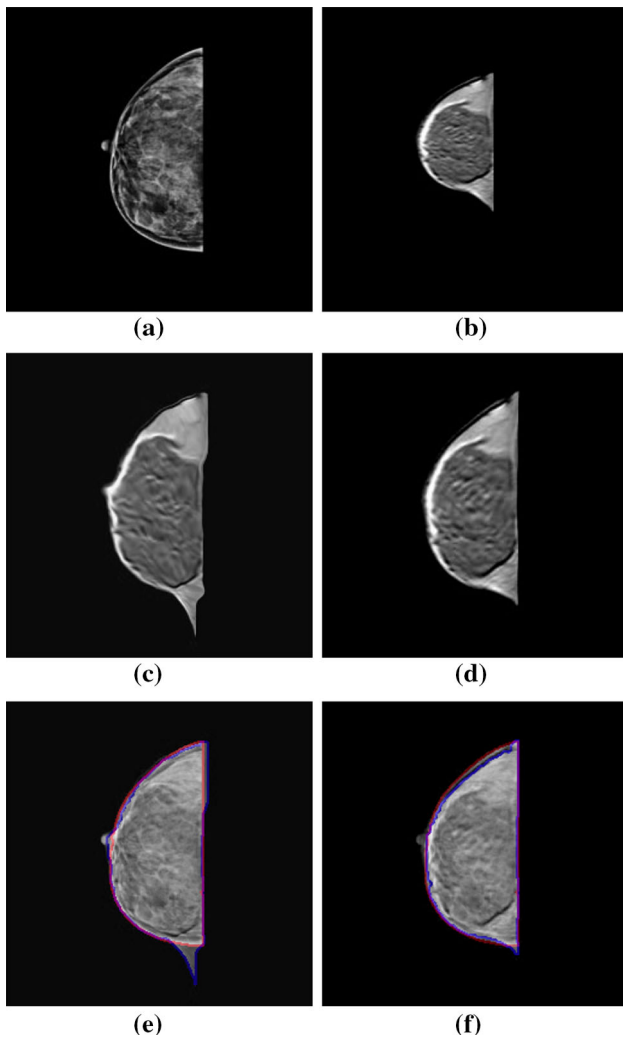
(a) **(b)**

(c) **(d)**

(e) **(f)**

Fig. 9 Examples for experiments of real breast images from the same patient. **a** Reference images of mammographic image, **b** floating image of MR modality. **c** The registration result from MI method. **d** The result from our method. **e** The overlay for **a** and **c** with the *contour lines*. **f** The overlay for **a** and **d** with the *contour lines*

Table 4 The mean metric values of different algorithms [mean value (95 % CI)]

Metrics	MI	FSR	*p* value
NMI	1.3716	1.4170	<0.05
	(1.3509,1.3924)	(1.3905,1.4436)	
NC	0.7859	0.8117	<0.05
	(0.7710, 0.8008)	(0.7945, 0.8289)	

The overlay (shown as Fig. 9f) of (a) and (d) illustrates that the contour lines from two images were properly superimposed.

The quantitative comparisons between our FSR method and the MI method in terms of NMI and NC are illustrated in Fig. 10. The *t* test results at the 0.05 level (given in Table 4) indicated that our method statistically outperformed the MI-based method.

All algorithms were implemented with C on Windows 7 operating system and performed on a DELL desktop with Intel(R) Core(TM) i7-4770 @ 3.4 GHz CPU. The average computation time for image structural representation using our FSR method was 27.5 s, and average computation time for image registration in the common space is 3 min and 48 s.

Discussion

In this paper, a new structural representation was constructed by fusing the structure tensor trace with local entropy to describe the geometric and structural properties of data. Through the fused structural representation, the multimodal data were converted into a new unified space that reflected its geometry uniformly across modalities, so that images in this new representation were matched using a simple L_2 distance as a similarity metric.
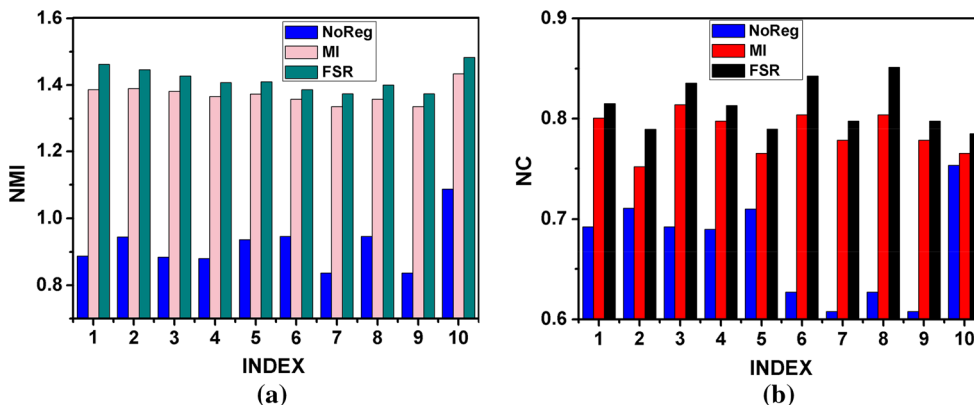

(a) **(b)**

Fig. 10 Quantitative comparison of registration results from MI method and FSR method in terms of **a** the normalized mutual information (NMI); **b** the normalized correlation (NC)

Experimental validation was performed on multimodality MRI brains images with artificial deformations and that on real images from different subjects. We compared our FSR method with other four methods: (1) the method with local entropy (LE); (2) the method using structure tensor (ST); (3) the method using spatially weighted local entropy (WLE); and (4) the conventional MI-based method.

A visual assessment of registration results is performed by observing the subregions marked with squares in Fig. 5. The region in blue square represents a subregion with rich structure information. Comparisons of those regions in the squares in Fig. 5 indicate that our proposed FSR algorithm outperforms other four algorithms, because the squared regions in the result image from the FSR method looks more alike the corresponding parts in the reference image. Such a result is plausible because structure information in these regions is highlighted by incorporating the structure tensor into the local entropy. Accuracy was quantitatively measured using the mean distance between the ground seeds and corresponding ones after registration, which were summarized in Table 1. From the table, we found that the mean errors obtained from FSR method were significantly smaller than the ones from LE method, WLE method and MI method, while they were marginally less than the ones using the ST method. The results indicated that using the fused structural representation, the registration algorithm could better correct the deformation between the reference image and the float image. Although our FSR method outperformed other four methods, we could see that structure tensor-based descriptor played a leading part in these two components. The same conclusion could also be obtained from Tables 2 and 3, where our FSR method achieved higher NC and NMI averages than other four methods. Both the structural tensor and the local entropy played the positive role.

Our method had some limitations in its current status. The registration was carried out at image level, and there was previous manual process for identifying the images with same local structure(s), and for instance, the corresponding multimodal MR images were all with tumor as shown in Fig. 4.

The experimental validation demonstrated that our method outperformed the other methods in comparison for the images with rich structural information, such as ventricle edges, the tumor regions in brain and the veins in breast images. However, the advantage of our proposed FSR method might attenuate if the images had relatively uniform intensity distribution and were in lack of the fine detail.

All experiments in this paper were performed on 2D images; however, our proposed method might be extended to 3D volumes registration by using the local entropy and structure tensor in the local hexahedron.

Conclusion

In this paper, a two-stage multimodal image registration algorithm was proposed for multimodal images. Images of different modalities were firstly converted to a unified common representation based on the structure tensor trace and local entropy. Experimental validation on multimodal MR brains images with artificial deformations and that on real multimodal brain MR images and breast images demonstrated that our proposed registration method outperformed the LE method, the ST method, the WLE method and the conventional MI-based method. Both the structure tensor and the local entropy played the positive role in the FSR.

Conflict of interest None.

References

1. Sotiras A, Davatzikos C, Paragios N (2013) Deformable medical image registration: a survey. IEEE Trans Med Imaging 32(7):1153–1190
2. Markelj P, Tomaževič D, Likar B, Pernuš F (2012) A review of 3D/2D registration methods for image-guided interventions. Med Image Anal 16(3):642–661
3. Roujol S, Benois-Pineau J, de Senneville BD, Ries M, Quesson B, Moonen CTW (2012) Robust real-time-constrained estimation of respiratory motion for interventional MRI on mobile organs. IEEE Trans Inf Technol Biomed 16(3):365–374
4. Otake Y, Armand M, Armiger RS, Kutzer MD, Basafa E, Kazanzides P, Taylor RH (2012) Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: incorporation of fiducial-based C-arm tracking and GPU-acceleration. IEEE Trans Med Imaging 31(4):948–962
5. Ferrant M, Nabavi A, Macq B, Black P, Jolesz F, Kikinis R, Warfield S (2012) Serial registration of intraoperative MR images of the brain. Med Image Anal 6(4):337–360
6. Wang X, Eberl S, Fulham M, Som S, Feng DD (2008) Data registration and fusion. In: Feng DD (ed) Biomedical information technology. Elsevier, Amsterdam, pp 187–210
7. Lu X, Zhao Y, Zhang B, Wu J, Li N, Jia W (2013) A non-rigid cardiac image registration method based on an optical flow model. Opt Int J Light Electron Opt 124(20):4266–4273
8. Lombaert H, Grady L, Pennec X, Ayache N, Cheriet F (2014) Spectral log-demons: diffeomorphic image registration with very large deformations. Int J Comput Vis 107(3):254–271
9. Zhang J, Wang J, Wang X, Feng D (2014) The adaptive FEM elastic model for medical image registration. Phys Med Biol 59:97–118
10. Zhuang X, Arridge S, Hawkes DJ, Ourselin S (2011) A nonrigid registration framework using spatially encoded mutual information and free-form deformations. IEEE Trans Med Imaging 30(10):1819–1828
11. Wachinger C, Navab N (2012) Entropy and Laplacian images: structural representations for multi-modal registration. Med Image Anal 16(1):1–17
12. Budde MD, Frank JA (2012) Examining brain microstructure using structure tensor analysis of histological sections. Neuroimage 63(1):1–10

13. Pluim JPW, Maintz JBA, Viergever MA (2000) Image registration by maximization of combined mutual information and gradient information. Med Image Comput Comput-Assist Interv MICCAI 2000 1935:452–461

14. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110(3):346–359

15. Yang L, Guo BL, Ni W (2008) Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. Neurocomputing 72(1–3):203–211

16. Zhang GY, Wittek A, Joldes GR, Jin X, Miller K (2014) A three-dimensional nonlinear meshfree algorithm for simulating mechanical responses of soft tissue. Eng Anal Bound Elem 42: 60–66

17. MICCAI (2012) Challenge on multimodal brain tumor segmentation. http://www.imm.dtu.dk/projects/BRATS2012. Accessed 20 Dec 2014

18. Bullitt E, Zeng D, Gerig G, Aylward S, Joshi S, Smith JK, Lin W, Ewend MG (2005) Vessel tortuosity and brain tumor malignancy: a blinded study. Acad Radiol 12:1232–1240

19. Marami B, Sirouspour S, Capson DW (2011) Model-based deformable registration of preoperative 3D to intraoperative low-resolution 3D and 2D sequences of MR images. In: Fichtinger G, Martel A, Peters T (eds) Medical Image Computing Computer-Assisted Intervention—MICCAI 2011, Vol 6891. Springer, Berlin, pp 460–467