

# LEARNING-BASED NON-RIGID IMAGE REGISTRATION USING PRIOR JOINT INTENSITY DISTRIBUTIONS WITH GRAPH-CUTS

Ronald W. K. So and Albert C. S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,  
Department of Computer Science and Engineering,  
The Hong Kong University of Science and Technology, Hong Kong  
{csswk,achung}@cse.ust.hk

## ABSTRACT

Non-rigid image registration is widely used in medical image analysis and processing. We recently proposed a novel learning-based similarity measure for non-rigid image registration. The novel similarity measure is constructed by using two Kullback-Leibler distances (KLD), which are based on the *a priori* knowledge of the joint intensity distribution of a pre-aligned image pair. In this paper, we propose a new formulation for the novel KLD based similarity measure such that it can be exploited in Markov random field (MRF) based non-rigid registration framework with the graph-cuts algorithm. We have compared the proposed formulation against two other similarity measures under the same MRF-based framework, and two state-of-the-art approaches. According to the experimental results, it is demonstrated that the proposed method can achieve high registration accuracy.

**Index Terms**— Non-rigid image registration, graph cuts, Kullback-Leibler distances.

## 1. INTRODUCTION

In medical image analysis and processing, non-rigid image registration is a widely used technique. Applications include deformation field recovery during surgery, and surgical procedure planning and evaluation [1]. The aim of non-rigid image registration is to find an optimal correspondence between pixels on the floating image  $I_f$  and the reference image  $I_r$ . Usually, the correspondence is represented by a transformation  $T$ . By using this optimal correspondence, i.e., the optimal transformation  $T^*$ , each pixel on the floating image can be mapped to a pixel on the reference image such that the value of a similarity measure can be minimized. In our previous work [2], we suggested to model non-rigid image registration as a multi-label assignment problem by using the Markov random field (MRF). The MRF-based energy function of the label assignment process is then minimized by graph-cuts via alpha-expansions [3] (It is called as graph-cuts based method

or GC hereafter). The formulation of the graph-cut method is,

$$\mathbf{T}^* = \arg \min_{\mathbf{T}} (E_d(I_f^{\mathbf{T}}(X), I_r(X)) + \lambda E_s(\mathbf{T})), \quad (1)$$

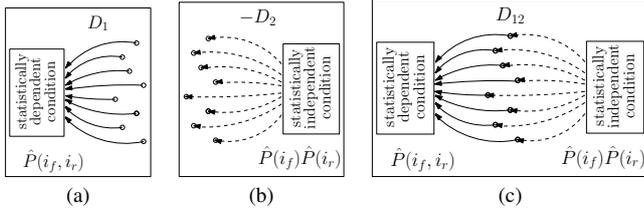
where  $E_d$ ,  $E_s$ ,  $I_f^{\mathbf{T}}$  and  $\lambda$  are data term, smoothness term, transformed floating image and the constant governing the smoothness. The data term represents the similarity measure and  $\mathbf{T}$  is the transformation which is in the form of deformation field. However, in the above graph-cuts based formulation, the data term is estimated locally based on the adjacent pixels within a neighborhood region. Therefore, some advanced similarity measures, e.g., mutual information (MI), etc., which are calculated globally, cannot be directly utilized in the graph-cuts based method. In this paper, we propose a new formulation for a novel learning-based similarity measure, which was originally proposed by Chung *et al.* [4] and estimated globally. With the new formulation, this measure can now be used in the graph-cuts based non-rigid registration method and it helps achieve high image registration accuracy.

## 2. KULLBACK-LEIBLER DISTANCES (KLD)

Recently, we proposed a new similarity measure for non-rigid image registration [5]. This similarity measure was used in rigid registration [4], and is based on *a priori* knowledge of the joint intensity distribution of a pre-aligned image pair. The measure is constructed by two KLDs which are denoted as  $D_1$  and  $D_2$ . Let  $P_o^T(i_f, i_r)$ ,  $\hat{P}(i_f, i_r)$ ,  $\hat{P}(i_f)$  and  $\hat{P}(i_r)$  be respectively the observed joint intensity distribution, expected joint intensity distribution, expected marginal distribution of floating image and expected marginal distribution of reference image, where  $i_f$  and  $i_r$  are the indices of the histogram bins of the floating and reference images respectively. Then, the formulations of  $D_1$  and  $D_2$  are,

$$D_1(P_o^{\mathbf{T}}(i_f, i_r) || \hat{P}(i_f, i_r)) = \iint P_o^{\mathbf{T}}(i_f, i_r) \log \frac{P_o^{\mathbf{T}}(i_f, i_r)}{\hat{P}(i_f, i_r)} di_f di_r, \quad (2)$$

$$D_2(P_o^{\mathbf{T}}(i_f, i_r) || \hat{P}(i_f, i_r)) = \iint P_o^{\mathbf{T}}(i_f, i_r) \log \frac{P_o^{\mathbf{T}}(i_f, i_r)}{\hat{P}(i_f) \hat{P}(i_r)} di_f di_r. \quad (3)$$



**Fig. 1.** Illustration of the functionalities of (a)  $D_1$ , (b)  $D_2$  and (c)  $D_{12}$ . The circle symbols represent the observed joint distributions of different image pairs before registration.

Figs. 1 (a) and (b) illustrate the functionalities of  $D_1$  and  $D_2$ . In the figure, the expected joint intensity distribution  $\hat{P}(i_f, i_r)$  and the product of expected marginal distributions  $\hat{P}(i_f)\hat{P}(i_r)$  represent the condition that the intensity values in the testing image pair are statistically dependent and independent respectively.  $D_1$  and  $D_2$  can be viewed as a pull force and a push force acting on the initial observed joint intensity distribution, which are illustrated by circles in Fig. 1. Minimizing  $D_1$  is equivalent to minimize the distance between the observed joint intensity distribution and the expected joint intensity distribution. Therefore, it pulls the observed joint intensity distribution towards the condition that the intensity values in the testing image pair are statistically dependent. On the other hand, maximizing  $D_2$  is the same as maximizing the distance between the observed joint intensity distribution and the product of the expected marginal distributions. Therefore, it can be viewed a push force which pushes the observed joint intensity distribution away from the condition that the image intensities in the testing pair are statistically independent. By combining the pull force and the push force (illustrated in Fig. 1 (c)), the final similarity measure  $D_{12}$  is defined as,

$$D_{12} = D_1 - D_2 = \iint P_o^T(i_f, i_r) \log \frac{\hat{P}(i_f)\hat{P}(i_r)}{\hat{P}(i_f, i_r)} di_f di_r. \quad (4)$$

### 3. GRAPH-CUTS BASED METHOD WITH $D_{12}$

In order to utilize  $D_{12}$  in the graph-cuts based method, we model the observed joint intensity distribution by using the Parzen windowing, i.e.,

$$P_o^T(i_f, i_r) = \frac{1}{|X|} \sum_{x \in X} g((i_f, i_r) - (I_f^T(x), I_r(x))),$$

where  $X$  and  $g$  denote the image domain and 2D Gaussian distribution respectively. Then Eq. (4) becomes

$$D_{12} = \sum_{x \in X} \frac{1}{|X|} \iint \log \left( \frac{\hat{P}(i_f)\hat{P}(i_r)}{\hat{P}(i_f, i_r)} \right) \cdot g((i_f, i_r) - (I_f^T(x), I_r(x))) di_f di_r. \quad (5)$$

Now, Eq. (5) is in the same form as the Eq. (7) in [6]. By following the procedure analogous to that discussed in [6], the similarity measure can be estimated by using Gaussian convolution. Eq. (5) can be further reduced as

$$D_{12} = \sum_{x \in X} \frac{-1}{|X|} \left( \log \left( \frac{\hat{P}(i_f)\hat{P}(i_r)}{\hat{P}(i_f, i_r)} \otimes g \right) \otimes g \right) (I_f^T(x), I_r(x)), \quad (6)$$

where  $\otimes g$  indicates Gaussian convolution. Now, Eq. (6) is our new formulation of  $D_{12}$  which is a summation among pixels in the image domain and can be utilized in the graph-cuts based method directly. Finally, our registration framework is defined as,

$$\mathbf{T}^* = \arg \min_{\mathbf{T}} (D_{12}(I_f^T(X), I_r(X)) + \lambda \mathbf{T}'), \quad (7)$$

where  $\mathbf{T}' = \sum_{(x,y) \in N} \|\mathbf{T}(x) - \mathbf{T}(y)\|$  is the smoothness term, for  $(x,y) \in N$  if  $y \in X$  is the adjacent pixel of  $x$ . This registration framework is then optimized by graph-cuts algorithm via alpha-expansion [3].

## 4. EXPERIMENTAL RESULTS

The quantitative experimental results of the proposed method (denoted as GC- $D_{12}$  hereafter) and the compared methods will be discussed in this section. We compared GC- $D_{12}$  against the graph-cuts based methods with two different similarity measures, sum of absolute difference and mutual information, which are denoted as GC-SAD [2] and GC-MI [7]. On the other hand, we also compared the proposed method with two state-of-the-art methods including Free Form Deformation (FFD) [8] and DEMONS [9]. For all the experiments, we assumed that the image pairs were affinely pre-registered and intensity normalized to the range [0 255]. The displacement window  $W$  of the graph-cuts based methods (including GC-SAD, GC-MI and GC- $D_{12}$ ) was set to  $\{0, \pm 1, \pm 2, \dots, \pm 15\}^2$ . In this setting, the deformation vectors assigned to pixels were chosen from a  $31 \times 31$  window. We modified the source code provided by Kolmogorov and Zabih [10] for the implementation of the graph-cuts algorithm. Implementations of FFD and DEMONS were adopted from the source code of the ITK library [11] in which  $15 \times 15$  control point grid was set for FFD. The 91<sup>st</sup> slices of the image volumes #04 and #20 of the Simulated Brain Database [12] were used as the reference image and the training image (for GC- $D_{12}$  only) respectively. Fig. 2 shows the reference image with its segmented images and the training image. We generated five floating images by applying five different artificial deformations to the reference image. The five floating images are named as Cases A, B, C, D and E, and showed in the left-most column of Fig. 3. All experiments were performed on an Intel 2.13 GHz dual-core CPU with 3 GB RAM.

The registration results of FFD, DEMONS, GC-SAD, GC-MI and GC- $D_{12}$  are respectively shown in the columns

Case	Tissue Class	Distributions of Absolute Intensity Errors (Mean $\pm$ SD)					
		Before Registration	FFD	DEMONS	GC-SAD	GC-MI	GC- $D_{12}$
Case A	WM	10.6479 $\pm$ 12.5717	3.8572 $\pm$ 4.1152	2.8604 $\pm$ 3.8747	0.4215 $\pm$ 1.9903	0.4556 $\pm$ 2.1131	0.2065 $\pm$ 0.8852
	GM	19.7615 $\pm$ 24.5911	4.7348 $\pm$ 5.3192	4.5281 $\pm$ 10.1937	0.4963 $\pm$ 2.3722	0.4541 $\pm$ 2.1660	0.2037 $\pm$ 0.9095
	CSF	21.6065 $\pm$ 22.6715	6.7096 $\pm$ 7.6762	6.5092 $\pm$ 11.7395	0.6051 $\pm$ 3.0877	0.6033 $\pm$ 2.9663	0.1984 $\pm$ 0.9944
	<b>Whole Image</b>	14.2316 $\pm$ 25.6177	4.3780 $\pm$ 6.1068	4.1457 $\pm$ 8.5574	0.2490 $\pm$ 1.7482	0.2118 $\pm$ 1.5318	0.1095 $\pm$ 0.6737
Case B	WM	2.2123 $\pm$ 6.8813	3.5560 $\pm$ 5.3306	1.3207 $\pm$ 1.7232	0.3748 $\pm$ 2.2254	0.1876 $\pm$ 1.1006	0.1173 $\pm$ 0.6973
	GM	4.8008 $\pm$ 10.4352	5.4411 $\pm$ 7.5203	1.3590 $\pm$ 2.3302	0.6927 $\pm$ 3.3107	0.2797 $\pm$ 1.4284	0.1634 $\pm$ 0.8461
	CSF	11.0332 $\pm$ 16.4650	10.7818 $\pm$ 12.6803	2.1143 $\pm$ 3.4919	1.4883 $\pm$ 5.3238	0.5407 $\pm$ 2.1647	0.2893 $\pm$ 1.1606
	<b>Whole Image</b>	4.3977 $\pm$ 13.8116	4.9032 $\pm$ 10.7390	1.6226 $\pm$ 4.8795	0.4680 $\pm$ 2.9115	0.2911 $\pm$ 2.0250	0.1215 $\pm$ 0.7608
Case C	WM	9.4930 $\pm$ 21.8914	5.6517 $\pm$ 12.4798	6.9019 $\pm$ 18.8221	0.5217 $\pm$ 2.0812	0.5094 $\pm$ 1.8529	0.2173 $\pm$ 0.8762
	GM	13.7174 $\pm$ 24.3150	7.2853 $\pm$ 11.7361	8.7060 $\pm$ 18.8371	0.5364 $\pm$ 2.4732	0.5201 $\pm$ 1.9356	0.2025 $\pm$ 0.9144
	CSF	14.1505 $\pm$ 24.0804	10.0187 $\pm$ 13.8591	9.1435 $\pm$ 17.7249	0.6261 $\pm$ 3.1866	0.4930 $\pm$ 1.8928	0.2007 $\pm$ 0.9837
	<b>Whole Image</b>	9.7191 $\pm$ 26.1838	5.5140 $\pm$ 13.0007	6.8966 $\pm$ 20.2661	0.3093 $\pm$ 1.8442	0.2652 $\pm$ 1.4335	0.1125 $\pm$ 0.7028
Case D	WM	9.6704 $\pm$ 10.7596	3.8462 $\pm$ 4.1204	2.5381 $\pm$ 3.0245	0.0884 $\pm$ 0.8623	0.0626 $\pm$ 0.5965	0.0410 $\pm$ 0.3975
	GM	14.9830 $\pm$ 16.0844	4.5487 $\pm$ 4.9767	1.9304 $\pm$ 2.7518	0.0968 $\pm$ 0.9653	0.0746 $\pm$ 0.7150	0.0422 $\pm$ 0.4005
	CSF	22.6887 $\pm$ 20.0158	6.3138 $\pm$ 6.5170	2.3448 $\pm$ 3.8415	0.1079 $\pm$ 1.0910	0.0655 $\pm$ 0.6781	0.0485 $\pm$ 0.4608
	<b>Whole Image</b>	10.9989 $\pm$ 18.7432	4.0331 $\pm$ 5.7967	2.0416 $\pm$ 5.9587	0.1146 $\pm$ 1.2744	0.0711 $\pm$ 0.7738	0.0483 $\pm$ 0.4566
Case E	WM	15.7858 $\pm$ 19.4334	7.3530 $\pm$ 9.7894	9.3675 $\pm$ 14.8211	0.5360 $\pm$ 2.3028	0.4509 $\pm$ 1.8172	0.1883 $\pm$ 0.8610
	GM	21.6458 $\pm$ 22.5526	9.4626 $\pm$ 11.3451	10.6319 $\pm$ 16.6659	0.6862 $\pm$ 2.9013	0.5679 $\pm$ 2.2531	0.2664 $\pm$ 1.0451
	CSF	33.6614 $\pm$ 26.2508	14.5973 $\pm$ 15.7412	13.5003 $\pm$ 20.3604	1.1614 $\pm$ 4.3055	0.9691 $\pm$ 3.4020	0.3916 $\pm$ 1.3199
	<b>Whole Image</b>	16.5570 $\pm$ 27.0167	7.0052 $\pm$ 11.3985	7.5937 $\pm$ 16.6385	0.5447 $\pm$ 2.8222	0.4090 $\pm$ 2.0275	0.1999 $\pm$ 0.9233
All	WM	9.5619 $\pm$ 14.3075	4.8528 $\pm$ 7.1671	4.5977 $\pm$ 8.4531	0.3885 $\pm$ 1.8924	0.3332 $\pm$ 1.4961	0.1541 $\pm$ 0.7434
	GM	14.9817 $\pm$ 19.5957	6.2945 $\pm$ 8.1795	5.4311 $\pm$ 10.1558	0.5017 $\pm$ 2.4045	0.3793 $\pm$ 1.6996	0.1756 $\pm$ 0.8231
	CSF	20.6280 $\pm$ 21.8967	9.6842 $\pm$ 11.2947	6.7224 $\pm$ 11.4316	0.7977 $\pm$ 3.3989	0.5343 $\pm$ 2.2208	0.2257 $\pm$ 0.9839
	<b>Whole Image</b>	11.1809 $\pm$ 22.2746	5.1667 $\pm$ 9.4083	4.4601 $\pm$ 11.2600	0.3371 $\pm$ 2.1201	0.2496 $\pm$ 1.5583	0.1183 $\pm$ 0.7034

**Table 1.** Distributions of the absolute intensity errors within different tissue regions after performing FFD, DEMONS, GC-SAD, GC-MI and GC-KLD. Errors are listed in the format, MEAN  $\pm$  SD.

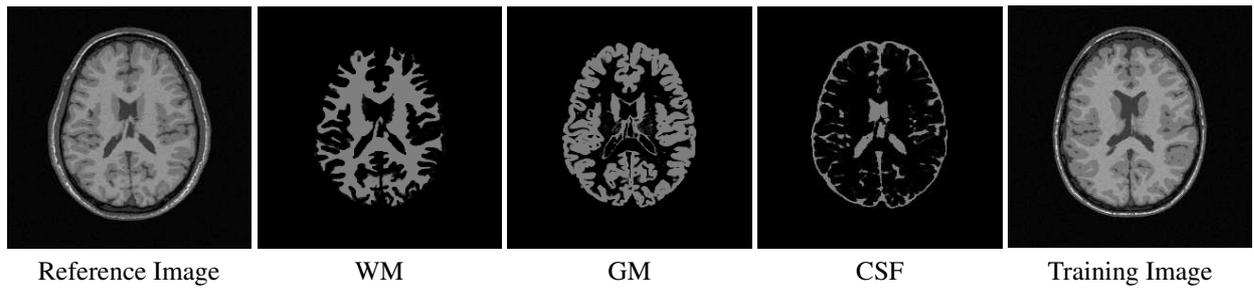
2-6 of Fig. 3. From the figure, it is observed that FFD cannot handle the cases with high frequency deformation like Cases B and E. This is because FFD applies hard constraints on the movements of the pixels in which only the control points can be moved freely. The displacements of other pixels are estimated by the interpolation among its neighborhood control points by a B-splines function. Therefore, registration including complicated deformations, like high frequency deformations in Cases B and E, are relatively difficult for FFD. On the other hand, DEMONS uses local intensity gradient to optimize the pixel movements. Although DEMONS allows high-degree of freedom, it may be trapped in local minima if some brain tissues of the reference image are overlapped with other tissues (including background) of the floating image or vice versa, like Cases C and E. This is because the pixels may be driven towards the wrong directions by DEMONS. Different from the FFD and DEMONS, the graph-cuts based methods optimize the registration process in a global manner by using the alpha-expansions which can provide high-degree of freedom. Therefore, graph-cuts based methods perform well in all the five cases. In order to evaluate the graph-cuts based methods, we have done a quantitative validation and the results are listed in Table 1. It demonstrates that the proposed method, GC- $D_{12}$ , outperforms other methods including GC-SAD and GC-MI which use the same registration framework. GC- $D_{12}$  can significantly reduce the averaged registration errors among different tissues as compared with GC-MI and GC-SAD (see the last row in Table 1). This shows that the improvement in registration accuracy mainly comes from the novel similarity measure  $D_{12}$ .

## 5. CONCLUSION

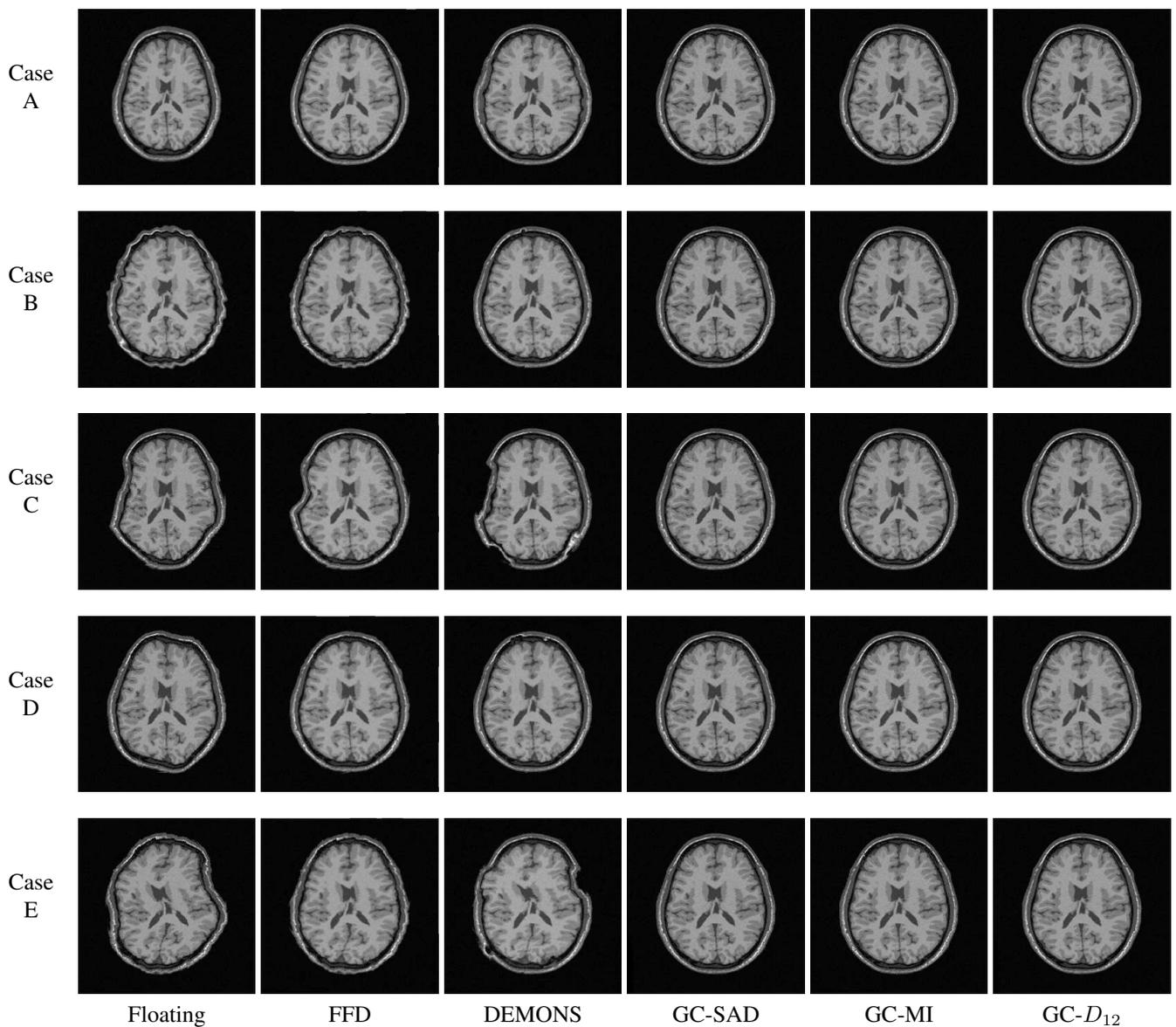
In this paper, we propose a new formulation for the novel learning-based similarity measure,  $D_{12}$ . By using this new formulation,  $D_{12}$  can now be exploited in the MRF-based non-rigid registration framework with graph-cuts optimization algorithm. The quantitative experimental results on five challenging registration cases demonstrate that the proposed method outperforms the other compared methods. Moreover, with the proposed method, the averaged absolute intensity errors of all tissues are reduced significantly.

## 6. REFERENCES

- [1] D. Rueckert, "Non-rigid registration: Techniques and applications," in *Medical Image Registration*, CRC Press, 2001.
- [2] T.W.H. Tang and A.C.S. Chung, "Non-rigid image registration using graph-cuts," *MICCAI LNCS*, vol. 4791, pp. 916-924, 2007.
- [3] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *PAMI*, vol. 23 (11), pp. 1222-1239, 2004.
- [4] Authors, "Publication of the authors under revision," .
- [5] R.W.K. So and A.C.S. Chung, "Multi-modal non-rigid image registration based on similarity and dissimilarity with the prior joint intensity distributions," *ISBI*, 2010.
- [6] J. Kim, V. Kolmogorov, and R. Zabih, "Visual correspondence using energy minimization and mutual information," in *ICCV*, 2003.
- [7] R.W.K. So and A.C.S. Chung, "Non-rigid image registration by using graph-cuts with mutual information," *ICIP*, 2010.
- [8] D. Rueckert, L. I. Sonoda, and et al, "Non-rigid registration using free-form deformations: Application to breast mr images," *TMI*, vol. 18 (8), pp. 712-721, 1999.
- [9] J. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *MedIA*, vol. 2 (3), pp. 243-260, 1998.
- [10] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *PAMI*, vol. 26 (2), pp. 147-159, 2005.
- [11] L. Ibanez and et al., *The ITK Software Guide*, 2003.
- [12] B. Aubert-Broche, A. Evans, and L. Collins, "A new improved version of the realistic digital brain phantom," *NeuroImage*, vol. 32 (1), pp. 138-145, 2006.



**Fig. 2.** Reference and its segmented images, and the training image obtained from Simulated Brain Database used in our experiments.



**Fig. 3.** Results of 5 different artificial deformation cases.