# Non-rigid Image Registration Using Graph-cuts

Tommy W.H. Tang 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
{cstommy,achung}@cse.ust.hk

**Abstract.** Non-rigid image registration is an ill-posed yet challenging problem due to its supernormal high degree of freedoms and inherent requirement of smoothness. Graph-cuts method is a powerful combinatorial optimization tool which has been successfully applied into image segmentation and stereo matching. Under some specific constraints, graph-cuts method yields either a global minimum or a local minimum in a strong sense. Thus, it is interesting to see the effects of using graph-cuts in non-rigid image registration. In this paper, we formulate non-rigid image registration as a discrete labeling problem. Each pixel in the source image is assigned a displacement label (which is a vector) indicating which position in the floating image it is spatially corresponding to. A smoothness constraint based on first derivative is used to penalize sharp changes in displacement labels across pixels. The whole system can be optimized by using the graph-cuts method via alpha-expansions. We compare 2D and 3D registration results of our method with two state-of-the-art approaches. It is found that our method is more robust to different challenging non-rigid registration cases with higher registration accuracy.

## 1 Introduction

Image registration is actively applied in the field of medical image analysis. Unlike rigid registration, non-rigid registration is an ill-posed problem due to its supernormal high degree of freedoms and inherent requirement of smoothness. Yet, there are a wide range of applications for non-rigid image registration [1].

The task of image registration is to find a transformation $T$ such that $I$ and $T(J)$ are spatially matched, according to an image-to-image dissimilarity measure, $C(I, T(J))$. $I$ and $J$ are referred as the source image and the floating image respectively and $T(J)$ refers to the resultant image after applying $T$ to $J$. Mathematically, the registration problem can be defined as finding the optimal transformation $T^*$ such that

$$T^* = \arg\min_T C(I, T(J)). \tag{1}$$

Unlike rigid image registration in which $T$ is restricted to a rigid transformation, for non-rigid image registration, there is still no common consensus in the literature regarding how the transformation $T$ should be modeled. Some models restrict $T$ to be of low degree of freedoms, such as affine, polyaffine [2] or

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part I, LNCS 4791, pp. 916–924, 2007.

control-points interpolated deformation [3] models. These models intrinsically constrain $T$ to be smooth or elastic and they are usually capable of representing an intra-patient deformation across time since there is a real physical underlying deformation between the images. However, in the case of inter-patient image registration, anatomical structures can vary significantly across patients both geometrically and topologically, a transformation of low degree of freedoms may not have the flexibility to represent these complex changes. Therefore, in principal, any hard constraints on the domain of $T$ should not be imposed. However, in Eqn. 1, we are optimizing $C(I, T(J))$ without posing any restrictions on $T$ and $T$ can map any points in $J$ to any points in $I$ without correlation across neighborhood pixels. Thus, $T$ needs regularization by adding a penalizing function $S(T)$ to penalize those $T$, which are not smooth. By modifying Eqn. 1, we get

$$T^* = \arg\min_{T} \quad C(I, T(J)) + \lambda S(T), \tag{2}$$

where $\lambda$ is a positive constant that controls the level of penalty for non-smooth $T$. If we consider $T$ as a displacement vector field, integrated magnitude of different derivatives is usually used as a criterion of smoothness in practice.

Two pioneer works of formulating non-rigid image registration, namely, Free-Form Deformations Based Method (denoted as **FFD** later) and Demons Based Method (denoted as **DEMONS** later), are widely used in the medical image analysis field and can be considered state-of-the-art. Rueckert *et al.* [3] proposed a method which modeled the local deformation by free-form deformation based on B-splines. In this method, only a regular grid of control points on the image are allowed to displace freely. The displacement of any other point is obtained from the displacements of its neighborhood control points, via B-spline interpolation functions. If a sparse set of control points is used, the transformation may not allow flexible movements of pixels to represent complicated deformation. We will show the effects in the experimental results section. Thirion [4] proposed a diffusion-based approach to non-rigid image registration. No hard constraints were imposed on the transformation $T$ so that each pixel can have its own displacement. In each iteration, the movement of any pixel in the floating image is based on its local intensity gradient and its intensity difference with the source image at the same position. It will naturally guarantee a decrease in (sum squared differences) SSD or (sum absolute differences) SAD by each iteration if the movement steps are sufficiently small. Since all pixels can move freely, a Gaussian smoothing step is applied at the end of each iteration in order to regularize the transformation. However, since the regularization is done after each iteration but not incorporated into the cost function, large displacements of pixels or sharp changes in the displacement field may not be penalized. Moreover, since the motions of pixels are highly depending on local intensity gradient, this method is highly sensitive to local artifacts. The effect will be demonstrated in the experimental results section.

In this paper, we formulate a new non-rigid image registration framework as a discrete labeling problem. Each pixel in the source image has a displacement label (which is a vector) indicating its corresponding position in the floating image, according to a similarity measure. A smoothness constraint based on first derivative is used to penalize sharp changes in displacement labels across pixels. The whole system can be optimized by using the graph-cuts method via alpha-expansions [5]. Through graph-cuts, the optimization process is not easily trapped in local minima and the solution is guaranteed to be within a known factor of the exact minimum. This makes the registration robustness and accuracy of our approach significantly better than other methods. To the best of our knowledge, our work is the first time where 3D labels are used in graph-cuts method, though using 2D labels in graph-cuts has been successfully applied in motion detection [5,6].

## 2 Theory and Methodology

### 2.1 Formulation of the Energy Function

Let $I$ and $J$ respectively be the source image and the floating image of dimension $d$ and $\mathbf{X}$ be the continuous spatial domain of both images. For any spatial point $\mathbf{x} = (x_1, x_2, ..., x_d) \in \mathbf{X}$, $I(\mathbf{x})$ and $J(\mathbf{x})$ are the intensity values (or feature vectors in general) at $\mathbf{x}$ of both images. In our formulation, a transformation $T$ is represented by a displacement vector field $\mathbf{D}$ that displaces every point $\mathbf{x}$ in $J$ away from its original position by the vector $\mathbf{D}(\mathbf{x}) \in \mathbb{R}^d$ to the new point $\mathbf{x} + \mathbf{D}(\mathbf{x})$. By modifying Eqn. 2, we can get

$$\mathbf{D}^* = \arg\min_{\mathbf{D}} \quad C(I(\mathbf{X}), J(\mathbf{X} + \mathbf{D})) + \lambda S(\mathbf{D}). \tag{3}$$

We use integrated absolute difference as the dissimilarity function $C$, and magnitide of first derivative terms as the smoothness function $S$. It yields

$$\mathbf{D}^* = \arg\min_{\mathbf{D}} \quad \int_{\mathbf{X}} \|I(\mathbf{x}) - J(\mathbf{x} + \mathbf{D}(\mathbf{x}))\| \, d\mathbf{X} + \lambda \sum_{i=1}^{d} \int_{\mathbf{X}} \|\mathbf{D}_{(x_i)}\| \, d\mathbf{X}, \tag{4}$$

where $\mathbf{D}_{(x_i)}$ is the first derivative of $\mathbf{D}$ along direction $x_i$ and the differential element $d\mathbf{X} = dx_1 dx_2 ... dx_d$. Since everything is in the continuous domain, $\mathbf{D}$ can have infinite degree of freedoms theoretically. Here, we introduce the first discretization step, by discretizing $\mathbf{X}$ into pixels. This is a natural discretizing step as images are usually acquired in a discretized form. By replacing all integrals by summations and derivatives by finite differences, Eqn. 4 becomes

$$\mathbf{D}^* = \arg\min_{\mathbf{D}} \quad \sum_{\mathbf{x} \in \mathbf{X}} \|I(\mathbf{x}) - J(\mathbf{x} + \mathbf{D}(\mathbf{x}))\| + \lambda \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{N}} \|\mathbf{D}(\mathbf{x}) - \mathbf{D}(\mathbf{y})\|, \tag{5}$$

where $(\mathbf{x}, \mathbf{y}) \in \mathcal{N}$ iff x and y are adjacent pixels. Note that at this stage, $\mathbf{D}(\mathbf{x}) \in \mathbb{R}^d$ is still not discretized. Therefore, $\mathbf{x} + \mathbf{D}(\mathbf{x})$ in Eqn. 5 can be any non-integer valued vector, and $J(\mathbf{x} + \mathbf{D}(\mathbf{x}))$ needs to be computed using an interpolation function. Also, when $\mathbf{x} + \mathbf{D}(\mathbf{x})$ is outside the image domain, a pre-assigned background intensity value can be used.

In principle, Eqn. 5 can be optimized by any iterative optimization tools instead of using the graph-cuts method. However, in practice, the degree of freedoms of $\mathbf{D}$ can be as high as a billion. First, it may cost huge amount of time for the optimization process. Second, since $\mathbf{D}$ has value in each pixel position, it is a requirement that the step size of updating $\mathbf{D}$ is sufficiently small in each iteration in order to ensure a smooth field. Not only adding an extra time cost, this makes the optimization process highly sensitive to local minima. Yet, Eqn. 5 is still not solvable by the graph-cuts method without modifications. It will be addressed shortly in the next subsection.

### 2.2 Optimization Via Graph-cuts

Unlike other general purpose techniques such as simulated annealing, which can be extremely slow in practice to find good optima of an energy function, the graph-cuts method yields either a global minimum or a strong local minimum in polynomial time, under some specific conditions. In general, the graph-cuts method is used to solve labeling problems by minimizing energy function $E_f$ in the following form [5,7],

$$E_f = \sum_{p \in \mathcal{P}} D_p(f_p) + \sum_{(p,q) \in \mathcal{N}} V_{p,q}(f_p, f_q). \tag{6}$$

In Eqn. 6, $\mathcal{P}$ is the set of pixels, $\mathcal{N} \subset \mathcal{P} \times \mathcal{P}$ is a neighborhood system defined on $\mathcal{P}$, $f : \mathcal{P} \to L$ is a labeling function where $L$ is a set of labels, $f_i \in L$ is the label of pixel $i$ in $f$. The term $D_p(f_p)$ measures the penalty of assigning label $f_p$ to pixel $p$ and the term $V_{p,q}(f_p, f_q)$ measures the penalty of assigning labels $f_p, f_q$ to the neighborhood pixels $p, q$ respectively. The two summations are usually referred as the data term and the smoothness term.

Comparing with the form of function solvable by the graph-cuts method in Eqn. 6, it is not difficult to observe that our current energy function in Eqn. 5 is already in that form if we consider a 4-connected neighborhood system $\mathcal{N}$, i.e., $(\mathbf{x}, \mathbf{y}) \in \mathcal{N}$ iff $\mathbf{x}, \mathbf{y}$ are adjacent pixels.

To convert our optimization to a labeling problem, $\mathbf{D}(\mathbf{x}) \in \mathbb{R}^d$ should be limited into a finite set. Here, we perform the second discretization step. Also acting as a restriction of how far a pixel can be displaced, a discretized window $\mathcal{W} = \{0, \pm s, \pm 2s, ..., \pm ws\}^d$ of dimension $d$ is chosen such that $\mathbf{D}(\mathbf{x}) \in \mathcal{W}$. Note that $\mathcal{W}$ is the discretization of the continuous dimension-$d$ region $[-ws, ws]^d$ with sampling period $s$ along all directions. Also, if $s < 1$, displacements with sub-pixel units can be considered. Now, by using $\mathcal{W}$ as the set of labels that every $\mathbf{D}(\mathbf{x})$ can be assigned, the optimization in Eqn. 5 can readily be solved by using graph-cuts via a sequence of alpha-expansion ($\alpha$-expansion) [5] moves.

Given the current labeling $f$ for the set of pixels $\mathcal{P}$ and a new label $\alpha$, an $\alpha$-expansion move means: For any pixel $p \in \mathcal{P}$, it is considered either keeping its current label $f_p$ or changing its label to $\alpha$ in the next labeling $f'$. Obviously, an $\alpha$-expansion move is a two-label problem, with label 0 meaning $f'_p = f_p$ and label 1 meaning $f'_p = \alpha$. Kolmogorov & Zabih [7] showed that the graph-cuts method can find the exact minimum of a two-label problem if every $V_{p,q}$ term in Eqn. 6 satisfies the following inequality.

$$V_{p,q}(0,0) + V_{p,q}(1,1) \leq V_{p,q}(0,1) + V_{p,q}(1,0). \tag{7}$$

We now show that any expansion move of our formulation satisfies Eqn. 7. Given a current labeling $f$ and two adjacent pixels $\mathbf{x}, \mathbf{y}$ with $f_{\mathbf{x}} = \beta$ and $f_{\mathbf{y}} = \gamma$, where $\beta, \gamma \in \mathcal{W}$, an expansion move of new label $\alpha \in \mathcal{W}$ is considered.

- $V_{\mathbf{x},\mathbf{y}}(0,0) = \|\beta - \gamma\|$ is the cost when both $\mathbf{x}, \mathbf{y}$ choose their old labels $\beta, \gamma$.
- $V_{\mathbf{x},\mathbf{y}}(1,1) = \|\alpha - \alpha\| = \mathbf{0}$ is the cost when both $\mathbf{x}, \mathbf{y}$ choose new label $\alpha$.
- $V_{\mathbf{x},\mathbf{y}}(0,1) = \|\beta - \alpha\|$ is the cost when $\mathbf{x}$ retains $\beta$ but $\mathbf{y}$ chooses new label $\alpha$.
- $V_{\mathbf{x},\mathbf{y}}(1,0) = \|\alpha - \gamma\|$ is the cost when $\mathbf{x}$ chooses new label $\alpha$ but $\mathbf{y}$ retains $\gamma$.

Since $\alpha, \beta, \gamma \in \mathcal{W} \subset \mathbb{R}^d$ and $\| \cdot \|$ is the L2-norm operator, by the triangle inequality, we have $\|\beta - \gamma\| \leq \|\beta - \alpha\| + \|\alpha - \gamma\|$ for any vectors $\alpha, \beta, \gamma$. Thus, the inequality in Eqn. 7 is satisfied for any adjacent pixels $\mathbf{x}, \mathbf{y}$ and each of our $\alpha$-expansion move is globally optimal. Boykov *et al.* [5] have further proved that, in such a case, the $\alpha$-expansion algorithm can finally converge to a local minimum, which is within a guaranteed factor of the exact minimum.

## 3 Experimental Results

In all experiments, it is assumed that all pairs of images are affinely pre-registered and the intensities of the images are normalized to be within 0 and 255. For **FFD** and **DEMONS**, we used the implementations obtained from ITK [8]. In **FFD**, we used a $15 \times 15$ control point grid. In our method, we used $\lambda = 0.05 \times 255$ and $\mathcal{W} = \{0, \pm 1, \pm 2, ..., \pm 15\}^2$ ($\{0, \pm 1, \pm 2, ...., \pm 12\}^3$ for 3D) so that displacement label of a pixel was chosen from a $31 \times 31$ window ($25 \times 25 \times 25$ for 3D) centered at that pixel. For the graph-cuts algorithm, we used the source codes provided by Kolmogorov & Zabih [7]. All MR and segmented data used in our experiments were obtained from BrainWeb [9]. Some slices are shown in Fig. 1.

**Registration Robustness.** The left of Fig. 1 shows an axial slice from an MR dataset which was used as the source image and four different artificial deformations (Case A-D) were applied to generate four floating images shown in the left column of Fig. 2. These artificial deformations can resemble different intra-subject and inter-subject mapping behaviors. The registration outputs of **DEMONS**, **FFD** and our method are shown in the last three columns in Fig. 2. From the registration output, it is obvious that **FFD** failed in Cases B and D and **DEMONS** failed in Cases C and D. All other cases were successful.

The failures of **FFD** in Cases B and D were situations where a transformation with low degree of freedoms cannot model a complicated or high-frequency deformation. **FFD** only allows control points to freely displace but restricts other pixels' displacements to be an interpolation of the displacements of neighborhood control points. As predicted, both **DEMONS** and our method are capable of

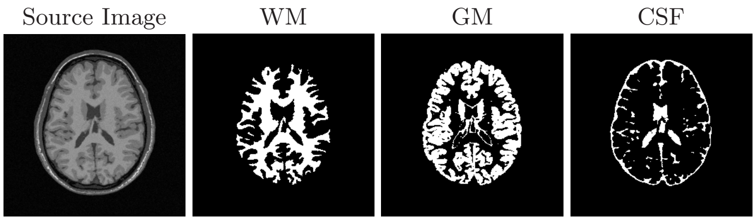


**Fig. 1.** Some original and segmented slices from BrainWeb used in our experiments

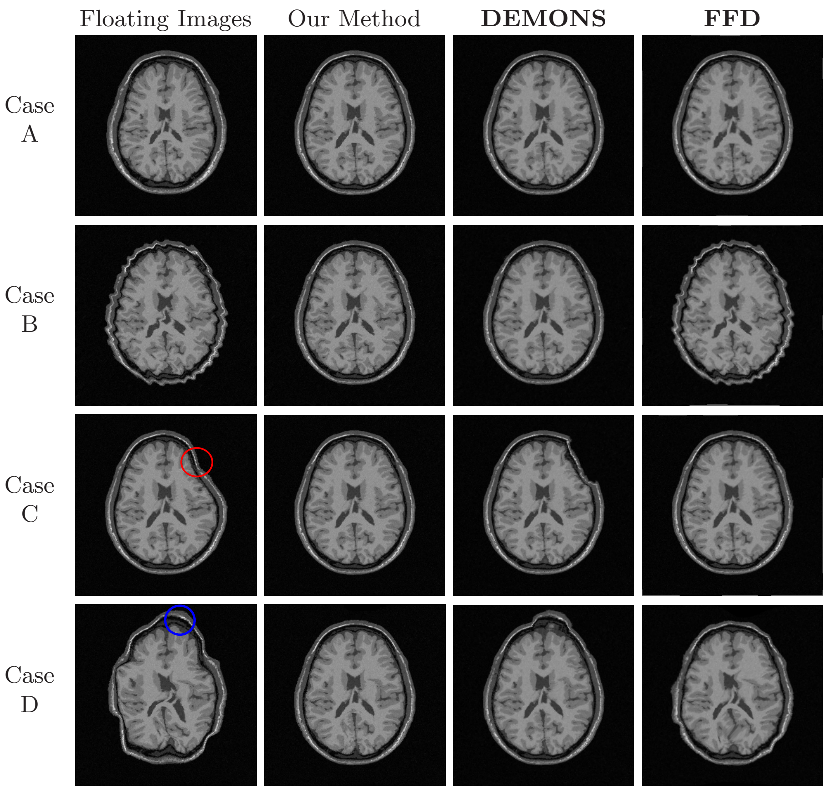


**Fig. 2.** *(Color Images)* Registration results of four different artificial deformation cases

restoring the images in such cases since no hard constraints are posed in the deformation models. The failures of **DEMONS** in Cases C and D were caused by local minima in the optimization process. The deformations in Cases C and D were considered large as some points are displaced more than 10 pixel-units. In Case C|D, before registration, some portions of the skull (circled red|blue in Fig. 2) in the floating image had its whole thickness being overlapping with the interior|background of the brain in the source image. Since **DEMONS** uses local intensity gradient to drive the movement of pixels, these initial overlapping

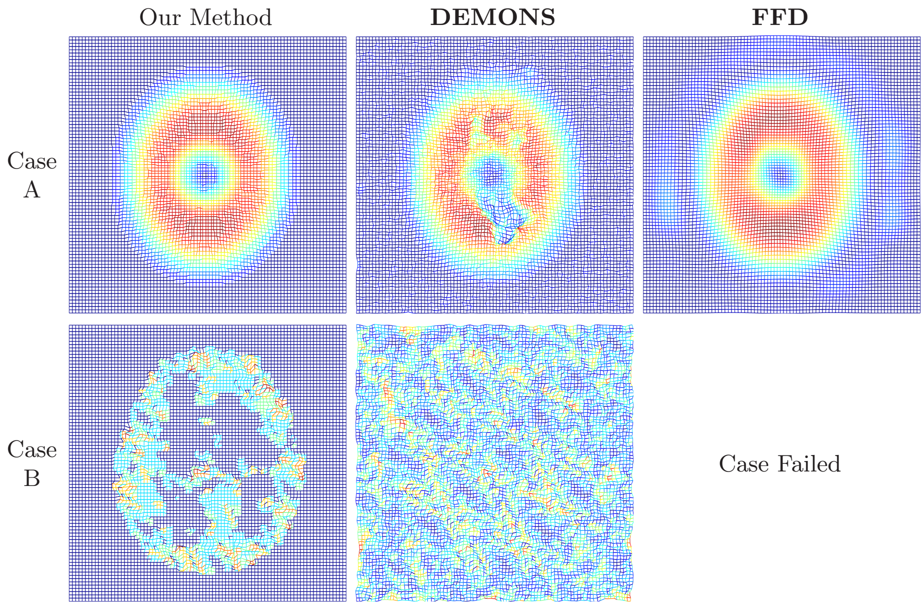


**Fig. 3.** *(Color Images)* Recovered deformation fields in Case A and Case B. Red color represents large displacements while blue color represents small displacements.

caused some pixels move towards the wrong directions and finally got trapped in local minima. Although our transformation model also has high degree of freedoms, our method still survive in this situation. It is because the graph-cuts method considers the labels of all pixels in a global manner in an $\alpha$-expansion move. Once the energy barrier is overcome, a group of pixels will together pursuit a large displacement in an $\alpha$-expansion step, without moving gradually through a series of small displacements.

**Smoothness of the Recovered Deformation Fields.** To compare the smoothness of the recovered deformation fields from different registration methods, we plot the fields for Cases A and B in Fig. 3. Case A corresponds to a squeeze-in deformation. As predicted, the fields recovered by **FFD** is the smoothest since **FFD** internally constrains the field to be a B-Spline interpolated transform. Comparing **DEMONS** and our method in both Cases A and B, it can be clearly seen that the field recovered by our method is much smoother than that recovered by **DEMONS**. It is because the smoothness constraint is kept globally in the cost function during the whole optimization process in our method.

**Registration Accuracy.** To evaluate for registration accuracy, we performed 10 (4 full and 6 downsampled) sets of 3D inter-patient registration by using **FFD**, **DEMONS** and our method. Fig. 4 shows the images and results of one of the sets using our method. Table 1 lists the distributions of pre/post-registration

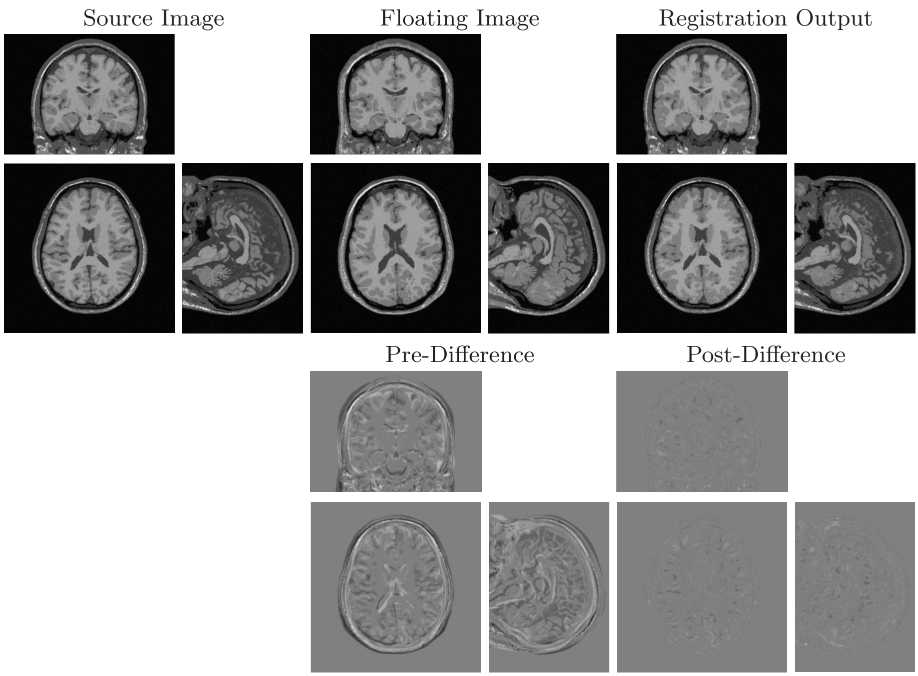


**Fig. 4.** 3D Registration results across two MR volumes. For each sub-figure, the middle slice across each axis of the volume is shown.

**Table 1.** Pre- and post-registration absolute intensity difference and tissue overlap measures of **FFD**, **DEMONS** and our method. Each number gives the mean value over all 3D registration tests.

| Tissue Class | Pre-Registration | **FFD** | **DEMONS** | Our Method |
|---|---|---|---|---|
| | Absolute Intensity Difference (Mean $\pm$ SD) | | | |
| Whole Image | $9.66 \pm 22.15$ | $7.31 \pm 17.17$ | $4.84 \pm 12.01$ | $2.65 \pm 6.30$ |
| | Tissue Overlap Measure | | | |
| WM | 0.4479 | 0.5041 | 0.6123 | 0.6754 |
| GM | 0.4575 | 0.5175 | 0.6282 | 0.7109 |
| CSF | 0.2395 | 0.3605 | 0.4721 | 0.5844 |

absolute intensity difference and the pre/post-registration overlap measures of three tissue classes, grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF). We adopted the overlap measure $\frac{\#(A \cap B)}{\#(A \cup B)}$, used by Crum *et al.* [10], where $A$ and $B$ denote the regions of the two images that belong to a specific tissue class. From the table, it is found that our method can consistently achieve higher registration accuracy than **FFD** and **DEMONS**.

## 4 Conclusion

We have proposed a new formulation to non-rigid image registration problem. First, we adopt a flexible deformation model, which allows every pixel to displace freely. This is essential for our method to recover any complicated deformation fields. Next, we present an energy function associated with the parameters of the deformation model, which are the displacement vector field $\mathbf{D}$. This function considers the dissimilarity measure of the images together with the smoothness requirement of the deformation field. Despite the supernormal high degree of freedoms in $\mathbf{D}$ as well as its smoothness requirement, we have successfully proved that our energy function can be globally optimized by using the graph-cuts method. The graph-cuts method also provides a solution with some degree of guarantee. Experimental results have demonstrated that our proposed method shows robustness to different challenging registration cases, e.g. large deformation, ripple distortion. It can be explained by the flexibility of our deformation model, as well as the power of the graph-cuts method to perform optimization in a global manner. Moreover, our method can achieve high registration accuracy.

## References


1. Rueckert, D.: Non-rigid registration: techniques and applications. In: Medical Image Registration, CRC Press (2001)
2. Arsigny, V., Pennec, X., Ayache, N.: Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. Medical Image Analysis 9(6), 507–523 (2005)
3. Rueckert, D., Sonoda, L.I., et al.: Non-rigid registration using free-form deformations: Application to breast MR images. TMI 18(8), 712–721 (1999)
4. Thirion, J.P.: Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis 2(3), 243–260 (1998)
5. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. PAMI 23(11), 1222–1239 (2001)
6. Winn, J., Jojic, N.: Locus: Learning object classes with unsupervised segmentation. In: ICCV, pp. 756–763 (2005)
7. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? PAMI 26(2), 147–159 (2004)
8. Ibanez, L., Schroeder, W., et al.: The ITK Software Guide, 1st edn. Kitware, Inc (2003)
9. Aubert-Broche, B., Evans, A., Collins, L.: A new improved version of the realistic digital brain phantom. NeuroImage 32(1), 138–145 (2006)
10. Crum, W.R., Rueckert, D., et al.: A framework for detailed objective comparison of non-rigid registration algorithms in neuroimaging. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3216, pp. 679–686. Springer, Heidelberg (2004)