Groupwise Registration via Graph Shrinkage on the Image Manifold

Shihui Ying[1,2] Guorong Wu[1] Qian Wang[1,3] Dinggang Shen[1]
[1] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, NC 27599, U.S.A
[2] Department of Mathematics, Shanghai University, Shanghai 200444, P.R. China
[3] Department of Computer Science, University of North Carolina at Chapel Hill, NC 27599, U.S.A
Yingshihui@gmail.com, grwu@med.unc.edu, qianwang@cs.unc.edu, dgshen@med.unc.edu

Abstract

Recently, groupwise registration has been investigated for simultaneous alignment of all images without selecting any individual image as the template, thus avoiding the potential bias in image registration. However, none of current groupwise registration method fully utilizes the image distribution to guide the registration. Thus, the registration performance usually suffers from large inter-subject variations across individual images. To solve this issue, we propose a novel groupwise registration algorithm for large population dataset, guided by the image distribution on the manifold. Specifically, we first use a graph to model the distribution of all image data sitting on the image manifold, with each node representing an image and each edge representing the geodesic pathway between two nodes (or images). Then, the procedure of warping all images to their population center turns to the dynamic shrinking of the graph nodes along their graph edges until all graph nodes become close to each other. Thus, the topology of image distribution on the image manifold is always preserved during the groupwise registration. More importantly, by modeling the distribution of all images via a graph, we can potentially reduce registration error since every time each image is warped only according to its nearby images with similar structures in the graph. We have evaluated our proposed groupwise registration method on both synthetic and real datasets, with comparison to the two state-of-the-art groupwise registration methods. All experimental results show that our proposed method achieves the best performance in terms of registration accuracy and robustness.

1. Introduction

In recent years, groupwise registration emerges as a new image normalization technique to simultaneously align a number of images to the latent population center. Since groupwise registration is able to avoid the bias in specifying reference image during registration, it has wide applications in both computer vision and medical imaging areas. For example, more reasonable appearance model of human face has been constructed in [2, 3] by jointly detecting the correspondences among 293 2D face images. Also, in many neuroscience studies, a large population of images is required to be normalized to the population center for better delineating the structural/functional difference due to brain development, aging and dementia [5, 14, 18].

Although a number of groupwise registration algorithms have been proposed [4–9, 11–13, 16, 17, 19], most of existing methods have the limitation of assuming only one center for a group of images, which prohibits further application on large complex population dataset. For example, studies on Alzheimer's disease usually need to register hundreds of brain images. However, due to considerable inter-subject variations, the group mean by simple average [12] is usually very fuzzy (due to the loss of anatomical details), especially when all images are far from being well registered. Then, taking the fuzzy group mean as the reference to guide the entire groupwise registration with clear individual images will eventually undermine the overall registration accuracy. More advanced weighting strategies, e.g., geometric mean [6] and sharp mean [17], have been proposed to address this issue, however, each image is independently registered to the population center without coordinating with its neighboring images on the image manifold. Intuitively, given the knowledge of image distribution (the locations of all image on a high-dimensional manifold), each image, even currently faraway from the population center, can better warp to the center through a deformation pathway which is relayed by a sequence of similar images, instead of being directly registered with the fuzzy group mean image.

In CVPR 2010, two groupwise registration papers [11, 19] proposed the similar idea to register all images by considering nearest neighboring images. However, only the topology of the local distribution at each image has been taken into account. Taking ABSORB algorithm [11] as example, each image is deformed locally w.r.t. its neighbors where all images are organized by a minimum spanning

tree. To guide moving all images towards the global center, the group mean image obtained by simple averaging is still used in ABSORB, because of the lack of knowledge of entire image distribution.

To deal with groupwise registration on large population dataset [5, 7], Wang *et al*. proposed a hierarchical registration framework to cluster images into a pyramid of classes [16]. Then, intra-class registration is performed to register all (similar) images within each class for generating the representative center image. The center images of all classes are further registered from the bottom to the top in the pyramid, until all images are registered. Therefore, this groupwise registration framework can efficiently register a large image data set with relatively better registration performance by addressing the challenges of registering two images with large anatomical differences.

Inspired by this ideal of hierarchical pyramid, we introduce the concept of graph into the groupwise registration, and then formulate the procedure of agglomerating all images into the common space as a dynamic evolution of graph shrinkage. Specifically, we consider each image as a node in the graph with each edge representing the geodesic pathway between the two nodes. Thus, the objective of our method becomes to shrink the graph by preserving its topology, until all nodes become close to each other. During graph shrinking, each image (or node in graph) is driven by the average velocities from its connected images (or nodes) in the graph, and warped along the geodesic on the image manifold. Since all images (or nodes) are warped simultaneously towards the center of graph, our method can achieve unbias in the registration.

Although ABSORB deforms each image locally w.r.t. its neighbors in the learned image manifold, our method has several advantages. (1) Our method is able to preserve the entire distribution of images, while ABSORB is only able to maintain the local distribution. (2) ABSORB still needs to estimate a tentative group-mean image for ensuring that each individual image is warped towards the group center, while our method completely avoids this step, thus introducing no bias. (3) ABSORB simply averages the deformations w.r.t. the neighboring images to warp each individual image, where the averaged deformation could be non-invertible. On the contrary, we address this limitation by employing diffeomorphism in our graph shrinkage procedure.

2. Methods

2.1. Background

Let $\{I_i\}_{i=1}^N$ be a group of N subject images to be registered. They are assumed to be sitting on the image manifold $\mathcal{M}$. In the groupwise registration method proposed by Joshi *et al*. [12], the population center is modeled as the Karcher mean I_c under an H^1 (Sobolev) metric, as described below:

$$I_c = \arg\min_{I \in \mathcal{M}} \sum_{i=1}^{N} d^2(I_i, I), \tag{1}$$

where $d(\cdot,\cdot)$ is the geodesic distance between two images on $\mathcal{M}$. For instance, d is defined as an H^1 metric in [12] and a metamorphosis metric in [6]. Eq. (1) describes that the group-mean image is a point on the manifold which has the close distance to all subject images (See Fig. 1.).

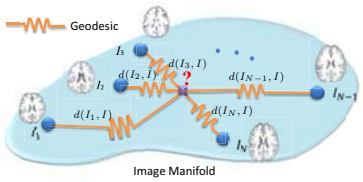


Figure 1. Karcher mean of N images on the image manifold.

For any pair of images $I_i, I_j \in \mathcal{M}$, their geodesic distance $d(I_i, I_j)$ is given as:

$$d(I_i, I_j) = \int_0^1 \|\dot{\gamma}_{ij}(s)\| ds = \int_0^1 \|v_{i,j}^s\| ds, \tag{2}$$

where $v_{i,j}^s$ is the velocity vector (or tangent vector in [10]) on the geodesic $\gamma_{ij}(s)$, with $\gamma_{ij}(0) = I_i$ and $\gamma_{ij}(1) = I_j$. The superscript s is the arc length parameter. According to the definition in [10], geodesic on manifold is a curve parameterized with constant velocity, i.e., $\|v_{i,j}^s\| = \|v_{i,j}^0\|, \forall s \in [0, 1]$. Hence, Eq. (2) turns to

$$d(I_i, I_j) = \|v_{i,j}^0\|, \tag{3}$$

by substituting $v_{i,j}^s$ with $v_{i,j}^0$ in Eq. (2). This indicates that the magnitude of the geodesic pathway from I_i to I_j eventually equals the length of the velocity vector at the initial point I_i $(s = 0)$. For convenience, in the following, we omit the superscript '0' in $v_{i,j}^0$ and use $v_{i,j}$ to represent the constant velocity vector of the geodesic from I_i to I_j.

2.2. Energy Function of Graph Shrinkage Model

The goal of groupwise registration is to simultaneously register all subject images to the population center. Here we assume the deforming of each individual image as a dynamic procedure of time variable t. Thus, $I_i(t)$ can be used to represent the deformed image I_i at time t. When t approach infinite (i.e., in the end of groupwise registration), the overall distance $F(t)$ between all pairs of the deformed images should be as small as possible:

$$F(t) = \sum_{i,j=1}^{N} d^2(I_i(t), I_j(t)) = \sum_{i,j=1}^{N} \|v_{i,j}(t)\|^2, \tag{4}$$

where $v_{i,j}(t)$ denotes the constant velocity vector from image $I_i(t)$ to image $I_j(t)$ at time t.

In general, deformable image registration will be performed to estimate each velocity vector $v_{i,j}(t)$ and further calculate the deformation pathway between $I_i(t)$ and $I_j(t)$ by $\exp(v_{i,j}(t))$, where exp is the exponential map [10]. Since it is challenging to register two images with large anatomical differences, minimizing $F(t)$ by considering the registration of all possible pairs of images might undermine the overall registration performance. Inspired by the ABSORB method, it is reasonable to consider the registration only between the two images with similar anatomical structures. Therefore, we introduce the variable e_{ij} to indicate whether images $I_i(t)$ and $I_j(t)$ are similar enough ($e_{ij} = 1$) or not ($e_{ij} = 0$). Thus, the weighted distance $F(t)$ in Eq. (4) can be rewritten as:

$$F(t) = \sum_{i,j=1}^{N} e_{ij} \|v_{i,j}(t)\|^2. \tag{5}$$

Next, we use a graph defined on the image manifold to interpret $F(t)$ in Eq. (5). Let $I(t) = \{I_i(t)\}_{i=1}^{N}$ be the graph node and $E = \{e_{ij} : i, j = 1, \cdots, N\}$ be the set of edges between two nodes in the graph. $e_{ij} = 1$ represents a link between $I_i(t)$ and $I_j(t)$. Otherwise, there is no direct link between $I_i(t)$ and $I_j(t)$ in the graph. Also, we define a weighted adjacency matrix $V(t) = (\exp(v_{i,j}(t)))$ as the velocity field from $I_i(t)$ to $I_j(t)$, which is used to describe the similarity of two images. Since we do not allow self-loop in the graph, we further define $e_{i,i}(t) = 0$ for all $i \in \{1, \cdots, N\}$. Then, the graph in our application can be defined as $\mathcal{G}(t) = (I(t), E, V(t))$. Since $\mathcal{G}(t)$ is an undirected graph, $V(t)$ is a symmetric matrix.

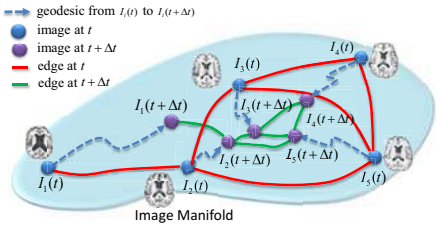


Figure 2. Demonstration of our proposed groupwise registration by graph shrinkage.

The principle behind $F(t)$ in Eq. (5) is demonstrated in Fig. 2. First, all images are assumed to be sitting in a high-dimensional manifold. Then, the topology of their distribution can be described by a graph, where the graph edges denote the local connections between graph nodes. Specifically, the velocity vector $v_{i,j}(t)$ is associated with each graph edge, where the integration along $v_{i,j}(t)$ forms the geodesic distance from $I_i(t)$ to $I_j(t)$. Thus, the minimization of $F(t)$ can be regarded as a dynamic graph shrinkage procedure, which deforms each image from $I_i(t)$ to $I_i(t + \Delta t)$ with the decreased overall geodesic distance, while keeping the topology of the entire graph. As shown in Fig. 2, all the images (in blue dots) at time t are the nodes in the graph where the graph edges are denoted by the red solid curves. At time $t + \Delta t$, each node $I_i(t)$ is deformed to its next position $I_i(t + \Delta t)$ (in purple dots) on the manifold. The graph $\mathcal{G}(t + \Delta t)$ (in purple dots and green curves) keeps the same topology with $\mathcal{G}(t)$ (in blue dots and red curves). As time t increases, all $I_i(t)$s are supposed to meet at the population center in the end of groupwise registration.

Obviously, the key steps in our method are (1) constructing the graph on the image manifold and (2) deforming image $I_i(t)$ toward the hidden population center at time t. We will explain these two steps next.

2.3. Graph Construction

In the beginning of groupwise registration ($t = 0$), an $N \times N$ distance matrix can be calculated with each element corresponding to the geodesic distance $d(I_i(0), I_j(0))$ between two images $I_i(0)$ and $I_j(0)$, where the geodesic distance can be estimated through the Log-Demons method [15]. The naïve solution for graph construction is to set a threshold h and then remove elements with geodesic distances higher than the threshold h, i.e.,

$$e_{ij} = \begin{cases} 1, & d(I_i(0), I_j(0)) < h \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Here, we go one step further to adaptively construct the graph according to the distribution of images. Specifically, two criteria are used to construct the graph: (C1) for any two nodes in the graph, there should be at least one path connecting these two nodes; (C2) the number of graph edges should be as low as possible, for saving the computational cost during the groupwise registration.

Accordingly, we propose a line-search-based method to determine the optimal threshold h. Given the geodesic distance between any two images $I_i(0)$ and $I_j(0)$, we set the search range within the low bound $b_L = 0$ and the upper bound $b_H = \max_{i,j} d(I_i(0), I_j(0))$. The optimal threshold h is updated as $h = b_L + \lambda(b_H - b_L)$, where $\lambda \in (0, 1)$ is a scalar specifying the step size in line search. Then, if the tentatively constructed graph satisfies the criterion (C1), the upper bound b_H will be decreased to h; otherwise, the low bound b_L will be increased to h. We repeat these steps until the low bound b_L meets the upper bound b_H. Given the optimal h, the graph can be constructed by Eq. (6) to guide the groupwise registration as described below.

2.4. Graph Shrinkage

As we formulate the problem of groupwise registration as the dynamic shrinkage of graph, it is critical to determine the deformation of each image $I_i(t)$ at time t, which

can consistently minimize $F(t)$ in Eq. (5) during the graph shrinking. To solve this minimization problem, we propose a descent method as detailed below.

Suppose each image has been deformed from $I_i(0)$ to $I_i(t)$ (at time t). It is natural to move $I_i(t)$ along the average velocity direction on the manifold according to its connected images in the graph $\mathcal{G}(t)$. Since the velocity vector sits on the tangent space of $I_i(t)$ on the manifold $\mathcal{M}$, it can be efficiently calculated by linear averaging as below:

$$\hat{v}_i(t) = \frac{1}{N_i}\sum_{j=1}^{N} e_{ij} v_{i,j}(t), \tag{7}$$

where $N_i = \sum_{j=1}^{N} e_{ij}$ is the number of connections for $I_i(t)$ in the graph. It is worth noting that we use Log-Demons [15] to estimate the velocity vector $v_{i,j}(t)$ between $I_i(t)$ and $I_j(t)$. The geodesic from $I_i(t)$ to $I_j(t)$ can be calculated by the exponential mapping from the vector space of stationary velocity field to diffeomorphism, i.e., $\exp(v_{i,j}(t))$.

Given a time increment Δt, image $I_i(t)$ is deformed to $I_i(t+\Delta t)$ along the geodesic $\exp(\hat{v}_i(t)\cdot\Delta t)$, where the velocity direction is steered by $\hat{v}_i(t)$ (Eq. (7)) and the step size is determined by Δt. We will explain the calculation of optimal Δt later. As shown in Fig. 3, the geodesics from $I_i(t)$ to $I_i(t+\Delta t)$, $I_i(t+\Delta t)$ to $I_j(t+\Delta t)$, $I_j(t+\Delta t)$ to $I_j(t)$, and $I_j(t)$ to $I_i(t)$ form a closed loop, given as:

$$\begin{aligned}\exp(v_{i,j}(t+\Delta t)) &= \exp(-\hat{v}_i(t)\cdot\Delta t)\circ\exp(v_{i,j}(t)) \\ &\quad\circ\exp(\hat{v}_j(t)\cdot\Delta t),\end{aligned} \tag{8}$$

where '$\circ$' denotes the composition of two geodesics[1]. If Δt is small enough, the velocity vector $v_{i,j}(t+\Delta t)$ can be approximated by applying Baker-Campbell-Hausdorff formula on Eq. (8) as:

$$\begin{aligned}v_{i,j}(t+\Delta t) &= -\hat{v}_i(t)\cdot\Delta t + v_{i,j}(t) \\ &\quad +\tau_{ji}(\hat{v}_j(t))\cdot\Delta t + o(\hat{v}_i(t)\cdot\Delta t),\end{aligned} \tag{9}$$

where $\tau_{ji} : T_{I_j}\mathcal{M} \mapsto T_{I_i}\mathcal{M}$ denotes the linearization of the left translation from $I_j(t)$ to $I_i(t)$, as shown in Fig. 3. Thanks to the left-invariant structure of the diffeomorphism group, we have the left invariant metric on this manifold. That is, $\langle\tau_{ji}(\hat{v}_j(t)), \tau_{ji}(\hat{v}_j(t))\rangle = \langle\hat{v}_j(t), \hat{v}_j(t)\rangle$.

Then, we have the following convergence theorem to prove that $F(t)$ is a monotonously decreasing function of time t along the velocity direction defined in Eq. (7).

Theorem 1 *The velocity fields defined in* (7) *make the objective function $F(t)$ strictly and monotonously decreasing.*

[1]Here, we follow the definition of composition in [1]

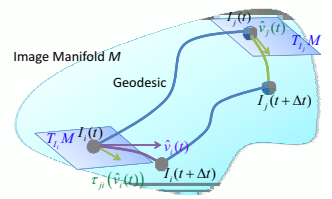


Figure 3. Demonstration of the translation of a vector, and relationship between the current and next images.

Proof: We directly calculate the derivative of the objective function (5) with respect to any fixed time variable t by

$$F'(t) = \lim_{\Delta t\to 0}\frac{\Delta F(t;\Delta t)}{\Delta t}, \tag{10}$$

where

$$\begin{aligned}\Delta F(t;\Delta t) &= F(t+\Delta t) - F(t) &(11)\\ &= \sum_{i,j=1}^{N} e_{ij}(\langle v_{i,j}(t+\Delta t), v_{i,j}(t+\Delta t)\rangle \\ &\quad -\langle v_{i,j}(t), v_{i,j}(t)\rangle). &(12)\end{aligned}$$

Substituting Eq. (9) to above equation, we have

$$\begin{aligned}\Delta F(t;\Delta t) &= \underbrace{-2\Delta t\sum_{i,j=1}^{N} e_{ij}\langle v_{i,j}(t), \hat{v}_i(t) - \tau_{ji}(\hat{v}_j(t))\rangle}_{\text{(I)}} \\ &\quad \underbrace{+\Delta t^2\sum_{i,j=1}^{N} e_{ij}\langle \hat{v}_i(t) - \tau_{ji}(\hat{v}_j(t)), \hat{v}_i(t) - \tau_{ji}(\hat{v}_j(t))\rangle}_{\text{(II)}} \\ &\quad +o(\Delta t^2\|\hat{v}_i(t)\|^2),\end{aligned} \tag{13}$$

where terms (I) and (II) are directly calculated with Eq. (7) as follows.

$$\begin{aligned}\text{(I)} &= \sum_{i=1}^{N}\langle\sum_{j=1}^{N} e_{ij} v_{i,j}(t), \hat{v}_i(t)\rangle - \\ &\quad \sum_{j=1}^{N}\langle\sum_{i=1}^{N} e_{ij}\tau_{ij}(v_{i,j}(t)), \hat{v}_j(t)\rangle &(14)\\ &= \sum_{i=1}^{N} N_i\|\hat{v}_i(t)\|^2 + \sum_{j=1}^{N}\langle\sum_{i=1}^{N} e_{ji} v_{j,i}(t), \hat{v}_j(t)\rangle &(15)\\ &= 2\sum_{i=1}^{N} N_i\|\hat{v}_i(t)\|^2, &(16)\end{aligned}$$

and

$$\begin{aligned}\text{(II)} &= \sum_{i,j=1}^{N} e_{ij}\|\hat{v}_i(t)\|^2 + \sum_{i,j=1}^{N} e_{ij}\|\hat{v}_j(t)\|^2 - \\ &\quad 2\sum_{i,j=1}^{N} e_{ij}\langle\hat{v}_i(t), \tau_{ji}(\hat{v}_j(t))\rangle &(17)\end{aligned}$$

$$= 2\sum_{i=1}^{N} N_i\|\hat{v}_i(t)\|^2 - 2\sum_{i=1}^{N}\langle\hat{v}_i(t), \sum_{j=1}^{N} e_{ij}\tau_{ji}(\hat{v}_j(t))\rangle \quad (18)$$

$$= 2\sum_{i=1}^{N}(N_i+1)\|\hat{v}_i(t)\|^2. \quad (19)$$

Therefore,

$$\Delta F(t;\Delta t) = -4\left(\sum_{i=1}^{N} N_i\|\hat{v}_i(t)\|^2\right)\cdot\Delta t + 2\left(\sum_{i=1}^{N}(N_i+1)\|\hat{v}_i(t)\|^2\right)\cdot\Delta t^2 + o(\Delta t^2\|\hat{v}_i(t)\|^2). \quad (20)$$

Obvious, $F'(t) = -4\left(\sum_{i=1}^{N} N_i\|\hat{v}_i(t)\|^2\right)$ is always negative for all $t > 0$. Therefore, the objective function (5) is always strictly and monotonously decreasing when t increases to infinity. With the lower bound of $F(t)$ being 0, the proposed algorithm is convergent. ■

2.5. Numerical Implementation

By Theorem 1, it is clear that the velocity vector defined by Eq. (7) makes the objective function (5) decreasing at any time t. To implement this minimization procedure, a numerical iterative process should be constructed. First, we discretize the continuous time t ($t \geq 0$) into $t^k (k = 1, 2, \cdots)$. Then, we calculate the velocity vectors $\hat{v}_i^k := \hat{v}_i(t^k), (i = 1, \cdots, N)$ by Eq. (7). Next, along these velocity vectors, a stepsize Δt^k should be determined. According to the convergent condition of the Taylor series of exponential map which makes the approximation (9) held, there should be $\Delta t^k \cdot \|\hat{v}_i^k\| < 1$ for all $i = 1, \cdots, N$, and therefore $\Delta t^k < 1/\max_i \|\hat{v}_i^k\|$. Under this condition, the increment $\Delta F(t^k;\Delta t) := \sum_{i,j=1}^{N} e_{ij}\left(\|v_{i,j}(t^k+\Delta t)\|^2 - \|v_{i,j}(t^k)\|^2\right)$ can be approximated by a positive definite quadratic function of Δt. Then, to accelerate the algorithm, Δt^k can be selected to make the increment $\Delta F(t^k;\Delta t)$ decreased as large as possible. Thus, it is straightforward to determine the optimal value of Δt by

$$\Delta t^k = \frac{\sum_{i=1}^{N} N_i\|\hat{v}_i^k\|^2}{\sum_{i=1}^{N}(N_i+1)\|\hat{v}_i^k\|^2}. \quad (21)$$

Especially, when the graph is fully connected, i.e. $N_i \equiv N-1$, $i = 1, \cdots, N$, we have

$$\Delta t^k \equiv \frac{N-1}{N}. \quad (22)$$

This is consistent with Theorem 3.1 in [13] which deforms one image w.r.t. all other images during groupwise registration. In general, the k^{th} stepsize Δt^k is selected by

$$\Delta t^k = \min\left\{\frac{1}{\max_i\|\hat{v}_i^k\|}, \frac{\sum_{i=1}^{N} N_i\|\hat{v}_i^k\|^2}{\sum_{i=1}^{N}(N_i+1)\|\hat{v}_i^k\|^2}\right\}. \quad (23)$$

2.6. Summary

In our groupwise registration method, we first use graph to model the distribution of all images on the manifold. Then, the groupwise registration is formulated as the graph shrinking procedure, where each image in the graph deforms along the graph edge on the manifold. Our whole method is summarized below.

Algorithm 1. Graph Shrinking Algorithm

Input: N affine aligned subjects $\{I_i\}_{i=1}^N$.
Output: The deformation pathway from I_i to the population center $\hat{I}_c$.
Initialization: Generate the graph section 2.3, given precision $\epsilon > 0$; Let $I_i^0 = I_i$, $i = 1, \cdots, N$ and set $k = 0$.
Repeat: (Graph shrinking)
  Do i
    Do j
      If $e_{ij} = 1$, estimate the velocity $v_{i,j}^k$ by the Log-Demons algorithm [15];
    End
    Compute $\hat{v}_i^k$ by (7), and estimate the step-size Δt^k by (23). Then, Update the warped image I_i^{k+1};
  End
  $k \Leftarrow k+1$;
Until: Convergence and output the deformation field from each I_i to the population center $\hat{I}_c$ by $\exp(\Delta t^k\hat{v}_i^k)\circ\cdots\circ\exp(\Delta t^0\hat{v}_i^0)$.

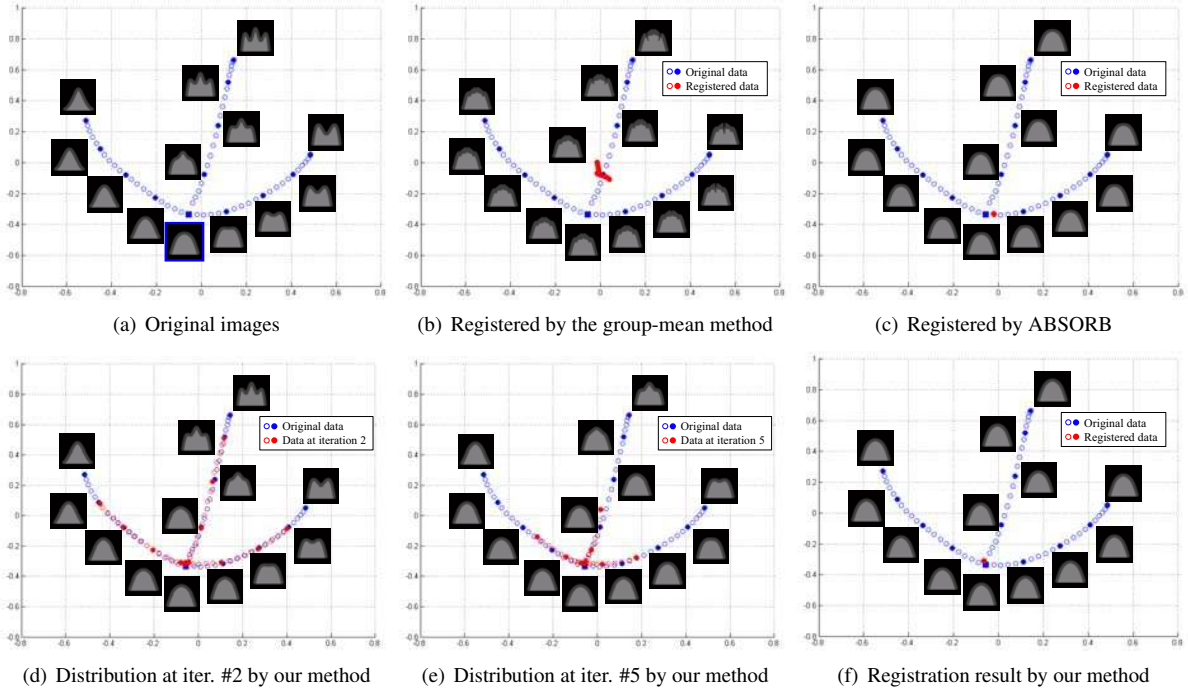
3. Experimental Results

In this section, we evaluate the registration performance of our proposed groupwise registration method on both synthetic and real datasets of infant brain images. For the sake of comparison, we also apply the group-mean registration method [12] and the ABSORB method [11] in our experiments[2].

3.1. Groupwise Registration on Synthetic Dataset

In this part, we first simulate the synthetic dataset that contains 61 2D images (sized 256×256) as follows. From the baseline image (indicated by the blue box in Fig. 4(a)), three different subsets of images are generated with each subset having 20 images, respectively. These three subsets have unique evolving patterns of folds. Typical images shown in Fig. 4(a) illustrate the 1, 2, and 3 folds in individual subsets. There are thus totally 61 ($20 \times 3 + 1$) images in the dataset. We project all images to a 2-dimensional space by performing PCA (Principal Component Analysis), and

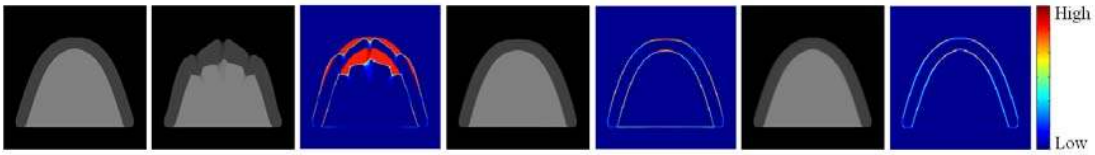
[2]The softwares of group-mean and ABSORB were downloaded at GLIRT (http://www.nitrc.org/projects/glirt/) and ABSORB (http://www.nitrc.org/projects/absorb/), respectively.



(a) Original images (b) Registered by the group-mean method (c) Registered by ABSORB

(d) Distribution at iter. #2 by our method (e) Distribution at iter. #5 by our method (f) Registration result by our method

Figure 4. The distributions of original images and registered images in the projected 2D space.



(a) (b) (c) (d) (e) (f) (g)

Figure 5. Ground truth and mean images by three method along with their residual errors. (a) is the ground truth, (b) (d) and (f) are the mean images by group-mean method, ABSORB and our method, respectively, and (c) (e) and (g) are the corresponding average residuals.

the distribution of images is shown in Fig. 4(a) where the solid circles indicate the accompanying images and the blue box indicates the baseline image, respectively.

After performing the groupwise registration with the group-mean method, ABSORB and our proposed method, we can obtain the final distributions of all registered images in the projected PCA space. The registered images are shown as red points in Fig. 4(b), (c) and (f) for the three methods, respectively. Fig. 4 (d) and (e) show the data distributions in the intermediate stages (i.e., after 2 iterations and 5 iterations) of our method. It is clear that the results of ABSORB and our proposed method are better than the group-mean method in that the estimated mean images in Fig. 4(c) and (f) are closer to the baseline image. It is also worth noting that the data distribution is gradually shrinked in Fig. 4(d)-(f), while the topology is consistently preserved. We compare the mean images in Fig. 5, with Fig. 5(a) showing the ground truth mean image. Fig. 5(b), (d) and (f) and Fig. 5 (c), (e) and (g) represent the estimated mean images and the average residual errors of all registered images by three methods, respectively. It is clear that our method achieves the best performance in terms of less registration errors.

3.2. Groupwise Registration on Longitudinal Infant Brains

The infant data used in this paper is a part of a large ongoing study of early brain development in normal children where T1-weighted MR brain images were collected by using a 3T SIEMENS scanner. 160 sagittal slices were obtained with parameters: TR=1900ms, TE=4.38ms, Flip Angle $= 7^{o}$. Data with motion artifacts was discarded and a re-scan was made when possible. Totally 7 longitudinal infant brains are used in this experiment, each with 6 time points (0, 3, 6, 9, 12 and 18 months of ages). The image size is $256 \times 256 \times 198$ and the voxel resolution is $1 \times 1 \times 1\text{mm}^3$.

Two longitudinal infant brains of two subjects at 6 different times are displayed in Fig. 6.

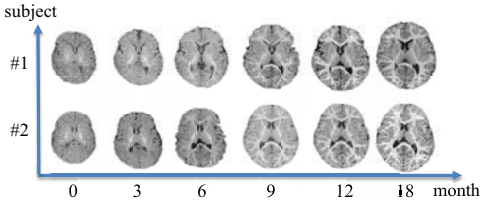


Figure 6. Representative longitudinal infant brains of 2 subjects at different times

The 3D renderings of the average images of all registered infant brain images by the conventional group-mean method, ABSORB, and our groupwise registration method are shown in Fig. 7. Due to large anatomical variations and particularly the dynamic intensity changes as shown in Fig. 6, it is difficult, especially in the beginning of groupwise registration, to find a representative group-mean image that well represents all infant brains. Therefore, the conventional group-mean method failed to align infant brains onto the common space, as indicated by very fuzzy average image shown in Fig. 7(a). Both ABSORB and our groupwise registration methods achieve more reasonable registration results than the convention group-mean method, since these two methods take the advantages of detected data manifold and consider registering only the similar images during the iterative registration procedure. On the other hand, by comparing the groupwise registration results of ABSORB in Fig. 7(b) and our method in Fig. 7(c), our method outperforms ABSORB in terms of the sharpness of average image and potentially the accuracy in registering each individual image.

To quantitatively evaluate the registration accuracy, we use the Dice ratio to measure the overlap degree between ROI A and ROI B, given as:

$$\mathrm{Dice}(A,B) = 2 \times \frac{|A \cap B|}{|A| + |B|}, \tag{24}$$

where $|\cdot|$ means the volume of the particular ROI. Since no template image is selected as a reference for groupwise registration, we will construct a labeled image in the common space by majority voting on ROIs of all registered images, in order to use the Dice ratio to evaluate the registration performance. Specifically, the Dice ratio of each ROI will be computed between the (constructed) labeled image and each registered individual image in the common space.

The Dice ratios on gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) by the conventional group-mean method, ABSORB, and our method are shown in Table 1. It is clear that our method achieves the highest Dice ratio compared to the other two groupwise registration methods, with overall 2.85% and 2.14% improvement, respectively. Moreover, the iterative evolution of the Dice ratio on GM, WM, and CSF during the groupwise registration is shown in Fig. 8, with green, blue and red curves representing the results by the conventional group-mean method, ABSORB, and our method, respectively. It can be observed that, as the number of iterations increases, the Dice ratios on all tissue types increase much more consistently in our method than in ABSORB method. The main reason is that ABSORB does not model the entire distribution of all images as we indicated before. Instead, in ABSORB, each individual image only considers its neighboring similar images when deforming to the estimated population center, which is not sufficient to preserve the topology of the entire image distribution during the groupwise registration. Therefore, it is difficult to guarantee that all the tentatively deformed images converge to the population center, which also explains why the curves of the Dice ratio by ABSORB is not smooth as shown in Fig. 8.

Table 1. Dice ratios of three brain tissues by three methods

	WM	GM	CSF	Overall
Group-mean method	71.28%	76.14%	57.71%	68.38%
ABSORB	72.51%	75.69%	59.05%	69.09%
Our method	**73.98%**	**79.31%**	**60.39%**	**71.23%**

4. Conclusion

In this paper, we have developed a novel groupwise registration by first introducing a concept of graph to model the entire image distribution. Then, the procedure of groupwise registration is formulated as the dynamic shrinkage of graph on the manifold, which brings the advantage of preserving the topology of the image distribution during the groupwise registration. Our proposed method has been evaluated on both synthetic data and real longitudinal infant brain data, where our method achieves the best registration result in comparison with the group-mean method and ABSORB.

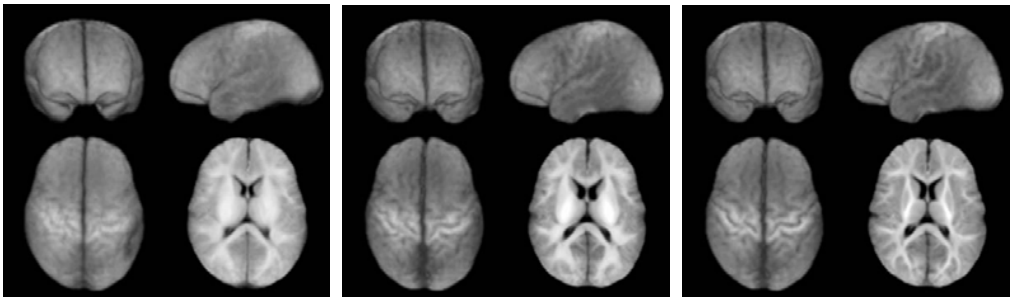
Acknowledgment

The research is supported partly by NIH grants EB006733, EB008374, EB009634, AG041721, NNSFC 61005002, 11101260, 60970149, and Discpline Project at the Corresponding Level of Shanghai (A.13-0101-12-005).

References

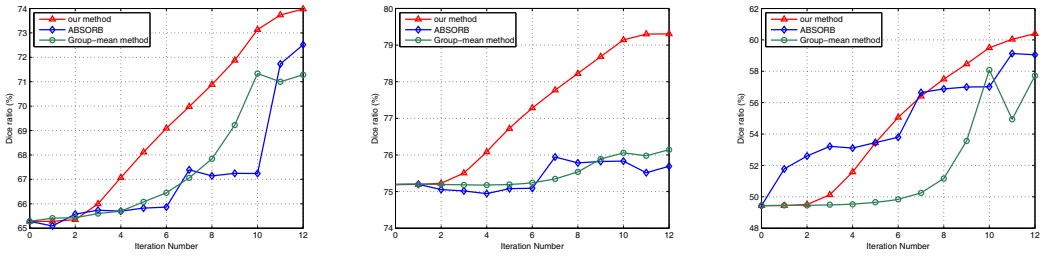
[1] J. Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113, 2007.

[2] T. Cootes, C. Twining, V. Petrovic, K. Babalola, and C. Taylor. Computing accurate correspondences across groups of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(11):1994–2005, 2010.



(a) Mean image by group-mean method (b) Mean image by ABSORB (c) Mean image by our method

Figure 7. Mean images by the three methods on longitudinal infant dataset.



(a) Evolution of Dice ratio on white matter (b) Evolution of Dice ratio on gray matter (c) Evolution of Dice ratio on CSF

Figure 8. Evolution of Dice ratios of three brain tissues during the groupwise registration by three methods on longitudinal infant dataset.

[3] D. Cristinacce and T. Cootes. Facial motion analysis using clustered shortest path tree registration. *Proc. MLVMA Workshop*, pages 1994–2005, 2008.

[4] B. Davis, P. Fletcher, and E. Bullit. Population shape regression from random design data. *Int. J. Comput. Vis.*, 90(2):255–266, 2010.

[5] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage*, 55(3):1073–1090, 2011.

[6] P. Fletcher, S. Venkatasubramanian, and S. Joshi. The geometric median on riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1-S1):S143–S152, 2009.

[7] V. Fonov, A. Evans, and K. Botteron. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1):313–327, 2011.

[8] S. Gerber, T. Tasdizen, P. Fletcher, S. Joshi, R. Whitaker, and ADNI. Manifold modeling for brain population analysis. *Med. Ima. Anal.*, 14:643–653, 2010.

[9] J. Hamm, C. Davatzikos, , and R. Verma. Efficient large deformation registration via geodesics on a learned manifold of images. *In: MICCAI*, pages 680–687, 2009.

[10] S. Helgason. *Differential Geometry, Lie Groups and Symmetric Space*, Academic Press, 2001.

[11] H. Jia, G. Wu, Q. Wang, and D. Shen. Absorb: atlas building by self-organized registration and bundling. *In: IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2785–2790, 2010.

[12] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23(S1):S151–S160, 2004.

[13] V. Noblet, C. Heinrich, F. Heitz, and J. Armspach. An efficient incremental strategy for constrained groupwise registration based on symmetric pairwise registration. *Pattern Recog. Lett.*, 33(3):283–290, 2012.

[14] A. Toga and P. Thompson. The role of image registration in brain mapping. *Ima. Vis. Comput.*, 19(1-2):3–24, 2001.

[15] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Symmetric log-domain diffeomorphic registration: a demons-based approach. *LNCS*, 5341:754–761, 2008.

[16] Q. Wang, L. Chen, P. Yap, G. Wu, and D. Shen. Groupwise registration based on hierarchical image clustering and atlas synthesis. *Human Brain Mapping*, 31:1128–1140, 2010.

[17] G. Wu, H. Jia, Q. Wang, and D. Shen. Sharpmean: groupwise registration guided by sharp mean image and tree-based registration. *NeuroImage*, 56(4):1968–1981, 2011.

[18] G. Wu, Q. Wang, D. Shen, and ADNI. Registration of longitudinal brain image sequences with implicit template and spatial-temporal heuristics. *NeuroImage*, 59(1):404–421, 2012.

[19] Y. Xie, J. Ho, and B. Vemuri. Image atlas construction via a intrinsic averaging on the manifold of images. *In: IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2933–2939, 2010.