# Joint Segmentation and Registration for Infant Brain Images

Guorong Wu[1], Li Wang[1], John Gilmore[2], Weili Lin[1], and Dinggang Shen[1(✉)]

[1] Department of Radiology and BRIC, University of North Carolina, Chapel Hill, NC 27599, USA
dgshen@med.unc.edu
[2] Department of Psychiatry, University of North Carolina, Chapel Hill, NC 27599, USA

**Abstract.** The first year of life is the most dynamic and perhaps the most critical phase of postnatal brain development. The ability to accurately characterize structure changes is very critical in early brain development studies, which highly relies on the performance of image segmentation and registration techniques. However, either infant image segmentation or registration, if deployed independently, encounters much more challenges than the adult brains due to dynamic appearance change with rapid brain development. Fortunately, image segmentation and registration of infant images can assist each other to overcome the above difficulties by harnessing the growth trajectories (temporal correspondences) learned from a large set of training subjects with complete longitudinal data. To this end, we propose a joint segmentation and registration algorithm for infant brain images. Promising segmentation and registration results have been achieved for infant brain MR images aged from 2-week-old to 1-year-old, indicating the applicability of our joint segmentation and registration method in early brain development studies.

## 1 Introduction

Human brain undergoes rapid physical growth and functional development from birth to 1 year old. The ability to accurately measure the structural changes from MR (Magnetic Resonance) images at this period is indispensable for shedding new light on the exploration of brain development and also the early detection of neurodevelopmental disorder. For example, infants with autism were found to have 5 %–10 % abnormal enlargement in total brain volume at early development stage [1].

However, the automatic image segmentation and registration tools for processing a large amount of infant brain MR images lag behind the demands from ongoing neuroscience/clinical studies. Both image segmentation and registration are challenged by: (1) the dynamic appearance changes of brain tissues from birth to 1 year old [1], and (2) the fast and spatially-varied developments of brain anatomy and size, especially in the first year of life [1]. Consequently, either segmentation or registration, if deployed independently, is difficult to handle the above challenges.


B. Menze et al. (Eds.): MCV 2014, LNCS 8848, pp. 13–21, 2014.
DOI: 10.1007/978-3-319-13972-2_2

Since the imaging-based study on early brain development becomes more and more popular, a sufficient number of longitudinal infant brain images have been collected in the past years. Many subjects with complete longitudinal images have been well segmented by either human interactions or automatic methods with multi-modality information [2]. As demonstrated by many literatures, appropriate joint segmentation and registration could significantly improve both of their performances [3]. Motivated by this, we aim to develop joint segmentation and registration method to overcome the above challenges. The leverage to achieve this goal is the availability of a large number of longitudinal infant data (scanned at 2 weeks, 3 months, 6 months, 9 months, and 1 year of age) and their respective segmentation results. It is worth noting that these valuable data are often ignored when performing segmentation/registration for the new infant images by current methods.

In this paper, we present an accurate and robust approach for joint segmentation and registration of any two given infant brain images by using the knowledge learned from the training infant subjects with the complete longitudinal data and segmentation results. Specifically, we first establish accurate temporal correspondences for each training subject with complete longitudinal images, in order to learn the subject-specific growth trajectories [4]. Then, to deal with the potential large age gap between two new infant images, we first segment them separately with a sparse patch-based level set approach that allows each patch in the new infant image to look for similar patches in the respective training subjects with similar age and further combine the labels of matched patches in the training subjects to provide a good initialization for the level set segmentation. Afterwards, deformations between two new infant images can be estimated by deformable registration upon their segmented images, thus avoiding the challenge of directly registering two new infant images with large appearance difference. The refined registration allows all matched patches of longitudinal training images (regardless of their ages) to improve the segmentation of each new infant image. By alternating these segmentation and registration steps, we can iteratively refine both segmentation and registration results of two new infant images. The advantages of our method include: (1) avoiding the direct registration between two new infant images with dynamic appearance changes, and (2) improving both segmentation and registration performances by fully using the available information from a number of training infant subjects with complete longitudinal data and their subject-specific growth models. In experiments, improved segmentation and registration results have been achieved for the infant images aged from 2-week-old to 1-year-old.

## 2 Methods

Our goal here is to register a moving infant brain image $M_{t_2}$ with a fixed infant brain image $F_{t_1}$ and also simultaneously determine tissue maps for $F_{t_1}$ and $M_{t_2}$, where $t_1$ and $t_2$ are two different ages, each of which could be as young as 2-week-old or as old as 1-year-old. Assume that we have $N$ training subjects $\boldsymbol{I}^s$ ($s = 1, \ldots, N$) with longitudinal data $\boldsymbol{I}^s = \left\{I_t^s | t = 1, \ldots, T_s\right\}$. For each image sequence $\boldsymbol{I}^s$, we can apply state-of-the-art 4D segmentation method [2] to segment 3D image at each time-point into WM (white matter), GM (gray matter), and CSF (cerebral-spinal fluid), which can be denoted as

$\boldsymbol{L}^s = \{L_t^s | t = 1, \ldots, T_s\}$. With some human inspection, we can regard these segmentation results of longitudinal training images as the ground truth.

Next, we first estimate the growth trajectory (Sect. 2.1) to determine temporal correspondences for each point in the longitudinal data $I_s$, as designated by the purple dash curves in Fig. 1. Second, we use sparse patch-based label fusion method to calculate the tissue probability maps for $F_{t_1}$ and $M_{t_2}$ separately, by using only the training images at the respective time-point as the atlas images. For example, the label fusion on $F_{t_1}$ selects only the image patches from training images $\left\{I_{t_1}^s | s = 1, \ldots, N\right\}$ at time-point $t_1$ (as shown in the dash pink box of Fig. 1) to form the dictionary. The obtained tissue probability maps can be used as a good initialization for level-set approach for tissue segmentation (Sect. 2.2). In this way, we can just register the two segmented images, thus avoiding the difficulty of directly registering the original two images with different appearances (Sect. 2.3). Given the spatial correspondences between $F_{t_1}$ and $M_{t_2}$, we can further improve the segmentation accuracy by augmenting the dictionary with additional image patches from training images at all other time-points, not simply the similar time-point(s) (Sect. 2.4). By alternating these segmentation and registration steps, we can iteratively refine both segmentation and registration results for $F_{t_1}$ and $M_{t_2}$.

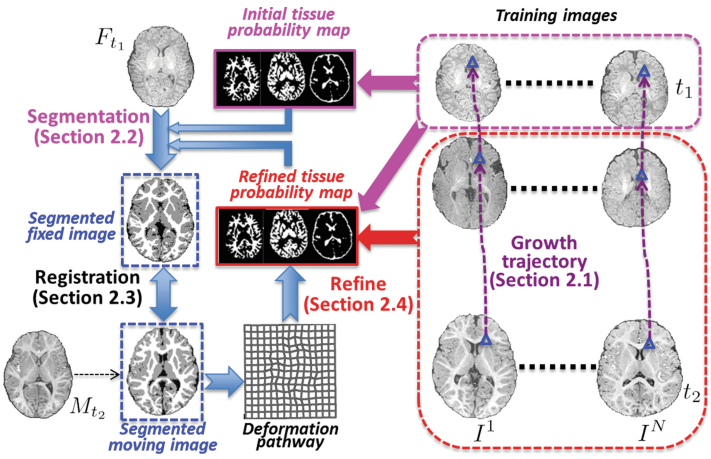


**Fig. 1.** The overview of joint segmentation and registration for infant brain images.

## 2.1 Learning Subject-Specific Growth Trajectories

For each training subject with complete longitudinal data $\boldsymbol{I}^s$ and segmentations $\boldsymbol{L}^s$, a conventional way is to independently register all follow-up images to the baseline image (first time-point). However, such independent image registration may tear down the coherence of temporal correspondences in each longitudinal data. Hence, we go one step further to apply a 4D image registration method [5] for jointly estimating the deformation fields $\mathbf{U}^s = \{u_t^s | t = 1, \ldots, T_s\}$ that can bring the image at each time-point to the latent common space. Thus, the temporal deformations $\varphi_{t \to t'}^s$ from time-point $t$ to

time-point $t'$ can be efficiently computed as the composition of inverse deformation $(u_t^s)^{-1}$ (from $I_t^s$ toward common space) and the forward deformation $u_{t'}^s$ (from common space to $I_{t'}^s$): $\varphi_{t\to t'}^s = \left(u_t^s\right)^{-1} \circ u_{t'}^s$, where $\circ$ denote for the deformation composition. In this way, we can use the estimated temporal deformation fields to form the growth trajectories of each training subject $\boldsymbol{I}^s$, as indicated by the purple dash curves in Fig. 1.

## 2.2 Sparse Patch-Based Level Set Segmentation

**Estimation of Tissue Probability Maps:** The initial tissue probability maps for each new infant image are very important to initialize the level set approach for achieving accurate tissue segmentation. Here, we use a patch-based label fusion method to estimate the tissue probability maps for each new infant image by considering the training images with same age as the atlases.

Let's take $F_{t_1}$ as example, where we consider all training images at $t_1$ time-point, $\{(I_{t_1}^s, L_{t_1}^s)|s = 1, \ldots, N\}$, as the atlases. We first affine register all atlases to $F_{t_1}$ and then apply deformable registration method to deform all atlases to $F_{t_1}$ image space. To determine the tissue type (WM, GM, or CSF) for each image point $x$ in $F_{t_1}$, we extract a referent patch $Q(x, t_1) \subset F_{t_1}$ centered at image point $x$. Next, we collect a number of atlas patches $P_s(v, t_1) \subset I_{t_1}^s$ across all training infant images $I_{t_1}^s$ at the same time-point $t_1$, with the center point $v$ sitting within a search neighborhood $n(x)$. Thus, all of these atlas patches form an over-complete dictionary $D(x, t_1) = \{P_s(v, t_1)|v \in n(x), s = 1, \ldots, N\}$. Since $F_{t_1}$ and $I_{t_1}^s$ are at the same time-point $t_1$, the appearances of these image patches are very similar. For clarity, we vectorize the reference patch $Q(x, t_1)$ into a column vector $b$. Also, we arrange each atlas patch $P_s(v, t_1)$ into a column vector $a_p$ and then assemble them into a matrix $\boldsymbol{A} = \left[a_p\right]_{p=1,\ldots,\eta}$, where $p = (v, s)$ is a bivariate index for the particular atlas patch $P_s(v, t_1)$ and $\eta = N \cdot |n(x)|$ denotes the total number of atlas patches.

Inspired by the power of sparse representation, we further look for a sparse coefficient vector $w$ to represent the reference patch $b$ by the dictionary matrix $\boldsymbol{A}$, i.e., $b \leftarrow \boldsymbol{A}w$, where each element in $w$ indicates the contribution of a particular atlas patch $a_p$ in representing the reference patch $b$. Thus, the estimation of $w$ falls to the classic LASSO (Least Absolute Shrinkage and Selection Operator) problem [6–8]:

$$\widehat{w} = \min_w \|b - \boldsymbol{A}w\|^2 + \lambda \|w\|_1, s.t. w > 0, \tag{1}$$

where $\lambda$ controls the sparsity of the coefficient vector $w$. Here, we specifically use $\widetilde{D}(x, t_1)$ to denote the set of selected image patches in $D(x, t_1)$ with the sparse coefficient $w_p > 0$. Since the tissue type for each $a_p$ is known, we can calculate the tissue probability w.r.t. WM, GM, and CSF, respectively. After repeating this procedure for every point in the fixed image $F_{t_1}$, we can obtain the tissue probability maps (as shown by a pink box in Fig. 1) to initialize the level set algorithm for segmenting $F_{t_1}$.

**Level Set Segmentation:** In level set algorithm, we employ three level sets, with their zero-level surfaces, respectively, denoting for interfaces of WM/GM, GM/CSF, and

CSF/background. The tissue probability maps can be integrated as prior knowledge into the coupled level set segmentation algorithm to improve the segmentation accuracy. Similarly, we can repeat the above procedure to segment the moving image $M_{t_2}$, except building the patch dictionary from the training images at time-point $t_2$. The tentative segmentation results of $F_{t_1}$ and $M_{t_2}$ are shown in the blue boxes of Fig. 1.

### 2.3 Symmetric Feature-Based Image Registration

Given segmentations for the fixed and moving images, we can deploy the state-of-the-art registration method, i.e., HAMMER [9], to register the two segmented images. Since geometric invariant moment (GMI) features are extracted from the segmented images, image registration is free of dynamic appearance changes in the original intensity images. Here, we further improve HAMMER by using the symmetric deformation estimation strategy, where we simultaneously estimate the deformation pathways $\phi_1$ (from fixed image) and $\phi_2$ (from moving image). It is worth noting that the deformed fixed image w.r.t. $\phi_1$ and the deformed moving image w.r.t. $\phi_2$ should be similar in the end of registration.

Since $\phi_1$ and $\phi_2$ are iteratively refined during registration, we use $k$ ($k = 0,.., K$) to denote the iteration. In the beginning of registration ($k = 0$), $F^{(0)} = F_{t_1}$ and $M^{(0)} = M_{t_2}$, along with the identity deformation pathways $\phi_1^{(0)}$ and $\phi_2^{(0)}$. In the following, we adopt the hierarchical deformation mechanism in HAMMER for establishing the correspondence between the deformed fixed image $F^{(k)} = F_{t_1}\left(\phi_1^{(k)}\right)$ and the deformed moving image $M^{(k)} = M_{t_2}\left(\phi_2^{(k)}\right)$, instead of always using $F^{(0)}$ and $M^{(0)}$ in HAMMER. Also, only a small number of key points with distinctive features are selected from $F^{(k)}$ and the $M^{(k)}$ to establish anatomical correspondences by matching the GMI features. The entire deformation pathways $\phi_1^{(k)}$ and $\phi_2^{(k)}$ are steered by the correspondences on these key points by requiring all other non-key points following the deformations on the nearby key points. With progress of registration, more and more key points are selected to refine the deformation pathways $\phi_1^{(k)}$ and $\phi_2^{(k)}$ regarding to $F^{(k)}$ and $M^{(k)}$, which is repeated until $F^{(k)}$ and $M^{(k)}$ become very similar in the end of registration. Finally, the deformation pathway from fixed image to moving image can be calculated by $\phi = \phi_1^K \circ (\phi_2^K)^{-1}$, where '$\circ$' denotes the composition of deformation pathway $\phi_1^K$ and the inverse deformation pathway $\left(\phi_2^K\right)^{-1}$.

### 2.4 Joint Segmentation and Registration by Using Growth Trajectories

Given the tentatively estimated deformations, we can further refine the tissue segmentation and then continue to improve the registration results with more accurate segmentation results. The key to achieve this goal is the augmented dictionary in tissue segmentation step since the refined registration will allow all matched image patches in other time-points to assist the segmentation, while these additional image patches are not included in the initial dictionary.

Taking $F_{t_1}$ as example, the initial dictionary used to segment each image point $x \in F_{t_1}$ uses only the image patches in time-point $t_1$ of training subjects (the dash pink box in Fig. 1). After registration, we assume the tentative corresponding location of $x$ in the moving image $M_{t_2}$ is $\phi(x)$. Since we know the growth trajectory in each training image, we can construct an augmented dictionary $D^*(x, t_1)$ for image point $x$, by now including: (1) *image patches at time-point* $t_1$: the training image patches from the initial dictionary $D(x, t_1)$ (at *time-point* $t_1$); (2) *image patches at time-point* $t_2$: the selected image patches in $\widetilde{D}(\phi(x), t_2)$ that is used to represent the moving image patch $Q(\phi(x), t_2)$, and (3) *image patches at all other time points*: all temporally corresponded image patches derived from $\widetilde{D}(\phi(x), t_2)$. Thus, we also know the appearance information at all other time points for each image patch in $\widetilde{D}(\phi(x), t_2)$ by traversing the growth trajectory learned in Sect. 2.1.

As shown in Fig. 1, in our joint segmentation and registration framework, segmenting infant image at a particular time-point can now utilize the information in the entire training images (pink and red boxes in Fig. 1). Since such additional subject-specific image patches are more pertinent to the image point $x$ than the training image patches from the entire population, the augmented dictionary can provide more useful information to guide tissue segmentation. By alternating the segmentation and registration steps, we can achieve state-of-the-art performance for both segmentation and registration for the infant brain images.

## 3 Experiments

In the training stage, we collect 24 infant subjects with complete longitudinal data as the training subjects, where each subject has T1- and T2-weighted images at 2 weeks, 3 months, 6 months, 9 months, and 12 months of age. All images were acquired from a Siemens head-only 3T MR scanner. T1-weighted images were acquired with 144 sagittal slices at resolution of $1 \times 1 \times 1\ mm^3$. T2-weighted images with 64 axis slices were obtained at resolution of $1.25 \times 1.25 \times 1.95\ mm^3$. For each subject, the T2-weighted image is aligned to the T1-weighted image at the same time-point and then further resampled to $1 \times 1 \times 1\ mm^3$. Note that we use T2-weighted images for segmenting 2-week-old and 3-month-old infant images, while use T1-weighted images to segment 6-, 9-, and 12-month-old infant images, considering the strong tissue contrast in the respective time-points for the respective MR images.

Since most of early brain development studies use the infant image at 1-year-old or older age as the atlas to discover structure development, we apply our joint segmentation and registration method to align 27 images at 2-week-old and 28 images at 3-month-old as moving images onto a 1-year-old infant image that is used as the fixed image (top right of Fig. 2). In the following, we evaluate the segmentation and registration performances one by one.

### 3.1 Evaluation of Segmentation Results

For each testing infant image, we have the manual segmentation results of WM, GM, and CSF. Here, we use Dice ratio to quantitatively measure the overlap degree between manual segmentations (used as ground truth) and our estimated tissue segmentations. The sparse patch-based level set segmentation algorithm (without using registration to refine) is used as the baseline method for comparison. The Dice ratios on each tissue type for 2-week-old and 3-month-old infant brain images are listed in Tables 1 and 2, respectively. After joint segmentation and registration, our method achieves overall 2.38 % and 1.95 % improvement in segmenting 2-week-old and 3-month-old infant brain images, respectively. Some typical improvements on 2-week-old infant brain image (top left of Fig. 2) are displayed in Fig. 2. It is clear that the initial mis-segmentations (in the blue boxes of Fig. 2) have been successfully corrected based on more and more accurate image registration (in the red boxes of Fig. 2).

**Table 1.** The Dice ratios of WM, GM, and CSF on 2-week-old infant brain images

| | WM | GM | CSF | Overall |
|---|---|---|---|---|
| Sparse-level set | 82.04 ± 5.17 | 81.02 ± 4.10 | 73.02 ± 5.77 | 78.69 ± 3.33 |
| Our method | 84.12 ± 2.38 | 83.53 ± 1.97 | 76.48 ± 2.98 | 81.07 ± 1.49 |

**Table 2.** The Dice ratios of WM, GM, and CSF on 3-month-old infant brain images

| | WM | GM | CSF | Overall |
|---|---|---|---|---|
| Sparse-level set | 82.53 ± 4.61 | 81.25 ± 3.53 | 75.14 ± 3.38 | 79.77 ± 2.68 |
| Our method | 84.35 ± 1.76 | 82.44 ± 2.11 | 77.00 ± 2.12 | 81.94 ± 1.27 |

### 3.2 Evaluation on Registration Results

A typical result of registering 2-week-old image to 1-year-old image is shown in the bottom right of Fig. 2. Since we have the manually labeled hippocampus region for both fixed image (1-year-old image as shown in the top right of Fig. 2) and eight 2-week-old infant brain images, we can further quantitatively evaluate the registration accuracy on 2-week-old images, by measuring the overlap degree between manual ground-truth and our estimated hippocampus (by deforming the hippocampal region of the fixed image to the image space of each individual 2-week-old image). In order to demonstrate the power of joint segmentation and registration, we apply HAMMER [9] on the baseline segmentations, obtained by using the sparse patch-based level set only (without refinement by joint registration). The average and standard deviation of Dice ratios on hippocampus is 70.13 ± 4.69 by HAMMER (based on the baseline segmentations) and 73.48 ± 2.05 by our joint segmentation and registration method, achieving almost 4.7 % increase in Dice ratio by our proposed method.

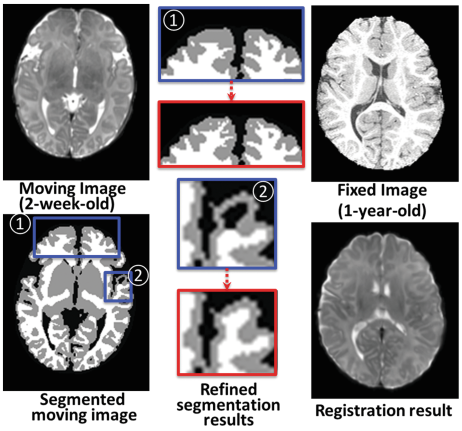


**Fig. 2.** Segmentation improvements on 2-week-old infant brain images, from results in blue boxes to the results in red boxes (Colour figure online).

## 4 Conclusion

In this paper, we propose a novel joint segmentation and registration method for infant brain images by using the growth trajectories learned from a large number of training subjects with complete longitudinal data. Specifically, image segmentation assists the registration by providing accurate tissue segmentations, which avoid the challenge of directly registering the two infant brain images with large appearance changes. In return, the refined image registration can bring more useful information to provide better tissue probability maps for guiding the level set based segmentation. Promising results for both segmentation and registration have been achieved, indicating the potential applicability of our method for early brain development study.

## References

1. Knickmeyer, R., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J., Hamer, R., Lin, W., Gerig, G., Gilmore, J.: A structural MRI study of human brain development from birth to 2 years. J. Neurosci. **28**, 12176–12182 (2008)
2. Wang, L., Shi, F., Yap, P.-T., Gilmore, J.H., Lin, W., Shen, D.: 4D Multi-Modality Tissue Segmentation of Serial Infant Images. PLoS ONE **7**, e44596 (2012)
3. Pohl, K., Fisher, J., Grimson, W.E.L., Kikinis, R., Wells, W.M.: A bayesian model for joint segmentation and registration. Neuroimage **31**, 228–239 (2006)
4. Miller, M.I.: Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. NeuroImage **23**, S19–S33 (2004)
5. Shen, D., Davatzikos, C.: Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping. NeuroImage **21**, 1508–1517 (2004)
6. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Stat. Methodol.) **58**, 267–288 (1996)

7. Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D.N., Zhou, X.S.: Towards robust and effective shape modeling: Sparse shape composition. Med. Image Anal. **16**, 265–277 (2012)
8. Zhang, S., Zhan, Y., Metaxas, D.N.: Deformable segmentation via sparse representation and dictionary learning. Med. Image Anal. **16**, 1385–1396 (2012)
9. Shen, D., Davatzikos, C.: HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans. Med. Imaging **21**, 1421–1439 (2002)

Springer

http://www.springer.com/978-3-319-13971-5

Medical Computer Vision: Algorithms for Big Data
International Workshop, MCV 2014, Held in Conjunction with MICCAI 2014, Cambridge, MA, USA, September 18, 2014, Revised Selected Papers
Menze, B.; Langs, G.; Montillo, A.; Kelm, M.; Müller, H.; Zhang, S.; Cai, W.(.; Metaxas, D. (Eds.)
2014, XI, 211 p. 78 illus., Softcover
ISBN: 978-3-319-13971-5