

# Spatial-temporal Constraint for Segmentation of Serial Infant Brain MR Images

Feng Shi<sup>1</sup>, Pew-Thian Yap<sup>1</sup>, John H. Gilmore<sup>2</sup>, Weili Lin<sup>3</sup>, and Dinggang Shen<sup>1</sup>

<sup>1</sup>IDEA Lab, <sup>3</sup>MRI Lab, Department of Radiology and BRIC,  
University of North Carolina at Chapel Hill, USA

<sup>2</sup>Department of Psychiatry, University of North Carolina at Chapel Hill, USA  
dgshen@med.unc.edu

**Abstract.** Longitudinal infant studies offer a unique opportunity for revealing the dynamics of rapid human brain development in the first year of life. To this end, it is important to develop tissue segmentation and registration techniques for facilitating the detection of global and local morphological changes of brain structures in an infant population. However, there are two inherent challenges involved in development of such techniques. *First*, the MR images of the isointense stage – the duration between infantile and early adult stages in the first year of life – have low gray-white matter contrast. *Second*, temporal consistency cannot be preserved if segmentation and registration are performed separately for different time-points. In this paper, we proposed a 4D joint registration and segmentation framework for serial infant brain MR images. Specifically, a spatial-temporal constraint is formulated to make optimal use of T1 and T2 images, as well as adaptively propagate prior probability maps among time-points. In this process, 4D registration is employed to determine anatomical correspondence across time-points, and also a multi-channel segmentation algorithm, guided by spatial-temporally constrained prior tissue probability maps, is applied to segment the T1 and T2 images simultaneously at each time-point. Registration and segmentation are iterated as an Expectation-Maximization (EM) process until convergence. The infant segmentations yielded by the proposed method show high agreement with the results given by a manual rater and outperform the results when no temporal information is considered.

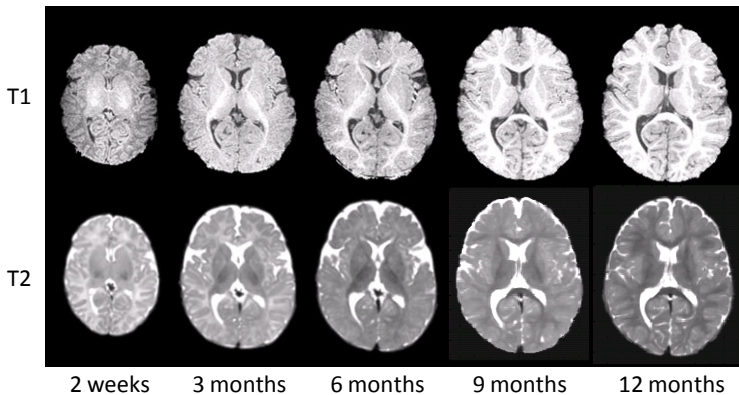
## 1 Introduction

Infants undergo rapid developments of brain structures and functions after birth, especially in the first year of life [1]. Tissue segmentation, a standard procedure in image analysis, is however challenging for infant images due to their inherently insufficient spatial resolution and low tissue contrast. Recently, numerous methods have been proposed for automatic brain segmentation of neonates (less than 3 months) [2-4] and infants (over 1-year-old) with varied success. However, little attention is paid to the segmentation of infants in the duration between birth and one year of age. Brain in this age range is affected by a combination of factors such as maturation and myelination. We show a series of longitudinal MR images in Fig. 1 for an infant scanned at an approximately 3-months interval, starting from two weeks.

We can observe that, in the T1 images, white matter (WM) initially has lower signal intensity than gray matter (GM) (e.g., 2 weeks), but eventually becomes higher and reaches the adult-like pattern after 9 months. A reversed contrast but same pattern can be observed in the T2 images. Our observations agree with Dietrich *et al.*'s findings [5], showing that three distinct gray-white matter signal intensity patterns appears in chronological order in images of developmentally normal infant – infantile (birth), isointense, and early adult (10 months onward). Images at isointense stage have the lowest gray-white matter contrast and have not been fully explored in previous studies.

To address this problem, we first make two observations. *First*, for the purpose of segmentation, T2 images are usually preferred in the infantile stage and T1 images in the early adult stage. This suggests that balance is needed between T1 and T2 images for guiding the segmentation. *Second*, infantile and early adult stage images can be segmented relatively easily compared to those from the isointense stage. This hints that priors could be constructed from both two stages and propagated to the isointense images for segmentation guidance.

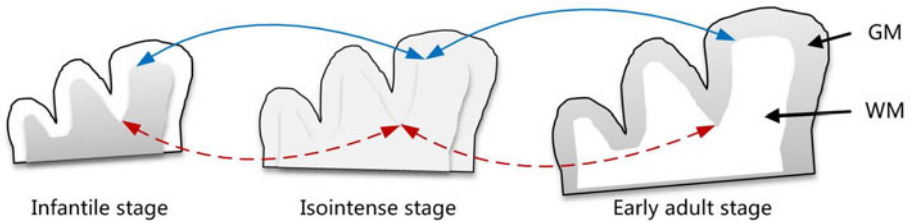
Temporal consistency is an important consideration factor in longitudinal infant studies. To analyze a 4D dataset, a conventional pipeline would perform segmentation on the images separately for each time-point. The segmented images are then registered across time-points for determining anatomical correspondences. Tissue density maps characterizing the volume and shape changes are finally employed for comparing brain changes over time. Great successes using this approach have been achieved in understanding early brain development [1]. However, inconsistency might be introduced by performing segmentation and registration independently with disregard of the temporal constraint, which could in the end reduce the reliability and statistical power of the findings.



**Fig. 1.** T1 and T2 images of an infant scanned at 2 weeks, 3 months, 6 months, 9 months, and 12 months. The gray-white matter contrast varies as a function of time, with three major stages – infantile, isointense, and early adult.

In this paper, we address the above-mentioned issues and propose a 4D joint registration and segmentation framework, the key idea of which is illustrated in Fig.2. Segmentation of images from the isointense stage is difficult, but this difficulty can be

ameliorated if we borrow information from neighboring time-points, providing constraint from temporal correspondences. By taking information from all time-points into consideration, temporal consistency can also be better preserved. To achieve this aim, we formulate a spatial-temporal constraint to make optimal use of the T1 and T2 images and to adaptively propagate the prior probability maps of each time-point to its neighboring time-points. In this process, 4D registration is employed to determine anatomical correspondences across time-points, and a multi-channel segmentation algorithm, which take into consideration the spatial-temporal constraint as prior, is used for segmenting the image at each time-point. Registration and segmentation are iterated as an Expectation-Maximization (EM) process until convergence. Details of the proposed method are given in the following sections. Note that although detection of within-tissue differences caused by myelination and maturation is also an important direction in early development study, this paper focuses on segmenting brain into general GM, WM, and CSF, in which WM contains both myelinated and unmyelinated tissue, GM contains both the cortical GM and basal ganglia.



**Fig. 2.** Spatial-temporal constraint for joint registration and segmentation of serial images from the first year of life. Three major gray-white matter patterns are shown. Probabilistic maps of tissues from the images of infantile and early adult stages are used to guide the segmentation of brain tissues in the isointense stage.

## 2 Method

We propose a 4D joint registration and segmentation framework, which alternates between 4D registration, spatial-temporal constraint construction (i.e., prior probability maps for each time-point), and multi-channel segmentation. T1 and T2 MR images are collected from 5 time-points – 2 weeks, 3 months, 6 months, 9 months, and 12 months. Images are skull-stripped with BSE [6], followed by manual editing to ensure the clean removal. Intensity inhomogeneity is corrected with N3 [7]. For initialization, we first segment the last time-point T1 image with an adaptive K-means algorithm. The resulting segmentation is taken as a longitudinal prior for guiding the initial segmentation of earlier time-point images independently, using the method described in [4]. This preliminary segmentation is fed as an input to our framework. In the following subsections, we will discuss the formulation of the spatial-temporal constraint, followed by the 4D joint registration and segmentation mechanism.

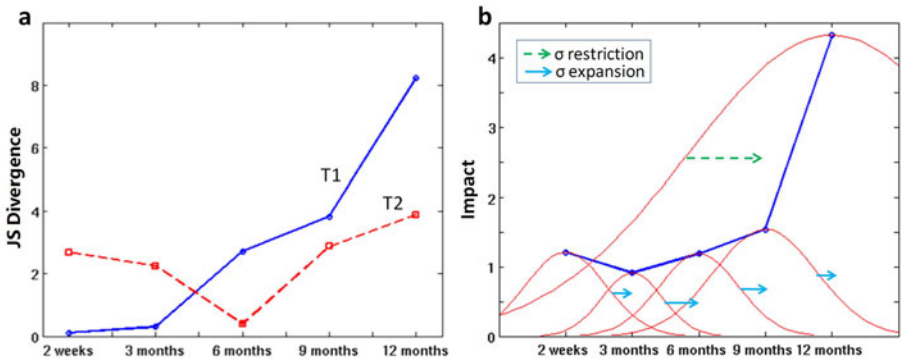
## 2.1 Formulation of Spatial-temporal Constraint

**T1-T2 Preference Factor.** As mentioned above, T2 is preferred for segmenting infantile images, and T1 for images from the early adult stage. To achieve balance and make optimal use of T1 and T2 images, we define *preference factors*  $\omega_{T_1}(t)$  and  $\omega_{T_2}(t)$  for each time-point  $t$ , where  $\omega_{T_1}(t) + \omega_{T_2}(t) = 1$ ,  $t = \{1,2,3,4,5\}$ , for reflecting the reliance of the segmentation algorithm on the T1 and T2 images, respectively. Since segmentation reliability is dependent on gray-white matter contrast, we compute the preference factors by comparing the histograms of GM and WM using the Jensen–Shannon (JS) divergence, which measures the difference of the two probability distributions and is symmetric. Initial GM and WM intensity distributions are obtained with the help of the initial segmented images and are updated as we have a more segmentation as we iterate through the registration-segmentation process. A plot of the JS values of the subject shown in Fig. 1 is given in Fig. 3(a). It can be observed that the tissue contrast of the T1 image is initially low and eventually increases to a value higher than that of the T2 image. The T2 image has a relatively good contrast throughout, except during the isointense stage. We thus define the *preference factors* as below:

$$\omega_{T_1}(t) = JS(t, T_{p1}) / (JS(t, T_1) + JS(t, T_2)) \quad (1)$$

$$\omega_{T_2}(t) = 1 - \omega_{T_1}(t) \quad (2)$$

where  $JS(t, T)$  is the JS divergence value, as a function of time-point  $t$  and image modality  $T = \{T_1, T_2\}$ . These preference factors will be employed in the multi-channel segmentation algorithm for utilizing the T1 and T2 images more effectively.



**Fig. 3.** (a) GM-WM JS divergence values signifying tissue contrasts in the T1 and T2 images of 5 different time-points. (b) Temporal impact for each time-point, and the Gaussian distribution of the impact factor for each time-point and their evaluation direction.

**Temporal Impact Factor.** We combine the T1 and T2 preference factors to obtain a measure called the *temporal impact factor*:

$$\begin{aligned}\omega(t) &= \omega_{T_1}(t) * JS(t, T_1) + \omega_{T_2}(t) * JS(t, T_2) \\ &= \frac{(JS(t, T_1))^2 + (JS(t, T_2))^2}{JS(t, T_1) + JS(t, T_2)}\end{aligned}\quad (3)$$

where the impact factor  $\omega(t)$  is weighted as a Contraharmonic mean of JS values of T1 and T2 images, which assigns more weight on the one with larger JS value.

This measure describes the across-time-point influence of an image. Time-point with high contrast image has a high impact value, implying that its influence will be propagated a long distance in the temporal dimension, and its probability map would have a greater impact on guiding the segmentation of other time-points. An example of this is the time-point of 12 months (see Fig. 3(b)). Due to the temporal changes in the brain, such as the generation of synapses and myelination, this impact is expected to be lower when propagated further along the temporal dimension. To simulate this, we incorporate the temporal impact factor into a Gaussian kernel and define:

$$G_{t,\sigma}(\tau) = \omega(t) \exp\left(-\frac{(\tau - t)^2}{2\sigma^2}\right) \quad (4)$$

where  $(\tau - t)$  is the temporal distance between the current time-point  $t$  and time-point  $\tau$  (with a unit temporal distance defined as equivalent to 3 months), and  $\sigma$  is a variance parameter for controlling the flatness of the Gaussian distribution. The parameter  $\sigma$  starts with  $\sigma = k\omega(t)$  and ends with a value of 1.  $k$  is a factor dealing with the scale difference of the temporal and the impact axes.  $G_{t,\sigma}(\tau)$  are indicated by red curves in Fig. 3(b).

The segmentation prior probabilistic map at a certain time-point  $t$ , is obtained via a linear combination of the segmentation probabilistic map at each time-point, weighted by the temporal impact factor:

$$p_t(L) = \frac{\sum_{\tau=1}^N G_{t,\sigma}(\tau) \times p_\tau(L|D)}{\sum_{\tau=1}^N G_{t,\sigma}(\tau)} \quad (5)$$

where  $L$  is a segmentation label,  $D$  is the deformation field pointing from time-point  $\tau$  to  $t$ , which is used to warp the probabilistic map from the space of time-point  $\tau$  to the space of time-point  $t$ , resulting in  $p_\tau(L|D)$ , and  $p_t(L)$  is the obtained prior probabilistic map of  $L$  at time-point  $t$ .

Notice that the segmentation result at each time-point is progressively refined and the dependence of other time-points is gradually reduced. Images from a certain time-point might have a large impact initially by providing significant prior information to other time-point images. As the segmentation result at the influenced time-point is getting better, less information needs to be borrowed from the neighboring time-points, and temporal impact factor of the neighboring time-point is decreased. On the other hand, an initially small temporal impact factor may also be increased, if the segmentation result of the current time-point is getting better, to help share the segmentation information with neighboring time-points. These two temporal impact

factor changing patterns can be controlled by the parameter  $\sigma$  in (4), to expand or to restrict as the temporal impact, as illustrated in Fig. 3(b).

## 2.2 4D Joint Registration and Segmentation Framework

4D registration helps determine anatomical correspondence between time-points, and also provides a means to warp the segmentation probabilistic maps of a certain time-point to other time-points. Segmentation, on the other hand, helps provides information for better registration of the images at each time-point. This can be described as an Expectation-Maximization (EM) process [8], with the E- and M-steps detailed as follows:

a) **E-step: Segmentation.** Given the intensity images  $\{I_{t,T}, T = T_1, T_2\}$  and prior probability maps  $p_t(L)$  obtained from M-step, at each time-point  $t$ , we can obtain a single segmentation  $S_t$  for both T1 and T2 images with multi-channel segmentation technique.

b) **M-step: Registration and Constrains Construction.** Deformation  $D$  is updated based on the new segmentations  $\{S_t\}$  from the E-step, providing anatomical correspondence between time-points. The parameter  $\sigma$  is changed to progressively control the temporal impact factor changing patterns. Then the spatial-temporal constraint  $p_t(L)$  is then computed by weighting the warped tissue probability maps  $p_t(L|D)$  across time-points, as detailed in subsection 2.1.

These two steps are iterated and progressively refine segmentation and registration. The program stops when segmentation results do not show further significant changes. In what follows, we briefly introduce the 4D registration and multi-channel segmentation approach used in this framework. More details can be found in [8, 9].

**4D Registration.** Given a group of segmentations  $\{S_t\}$  for the input image sets  $\{I_{t,T}, T = T_1, T_2\}$ , we need to determine the anatomical correspondences across time-points by a series of deformation fields  $D$ . A template set  $\{S_t^{Template}\}$  is first generated by duplicating the last time-point ( $t = 5$ ) segmented image (with best contrast) to each other time-point. To register the image sets  $\{S_t\}$  to  $\{S_t^{Template}\}$ , we hierarchically determine the optimal transformation which minimizes the difference between the attribute vectors of images in  $\{S_t\}$  and  $\{S_t^{Template}\}$  [9]. The attribute vector is composed of the Geometric Moment Invariants (GMI) for extraction of structural information. The cost function for the 4D registrations is:

$$E = E_F + E_B + E_S \quad (6)$$

where  $E_F$  and  $E_B$  are the attribute vector similarity cost functions related to the forward and backward transforms, and  $E_S$  is a spatial-temporal smoothness constraint. 4D registration has been evaluated and utilized in many applications such as in the studies of longitudinal changes of hippocampal volumes [9] and myocardial motion estimation [10].

**Multi-Channel Segmentation.** Having both T1 and T2 images allows us to make optimal use of the information from both modalities for better segmentation. Some multi-channel segmentation algorithms produce two individual results for T1 and T2

images. However, since both T1 and T2 images reflect the same tissue structures, producing a single segmentation result might be more appropriate. To this end, we employ a multi-channel segmentation algorithm which works with joint probability distributions. Denoting  $I$  as the intensity image, we view the T1 and T2 images as a single vector image with 2 elements at each voxel. We extend the single-channel Gaussian mixture model (GMM) segmentation [8] to its multi-channel counterpart by using axis-aligned multivariate Gaussians. In particular, for a voxel with vector  $X = (x_1, x_2)^T$ , we can estimate the mean  $\mu_i = (\mu_{i,1}, \mu_{i,2})^T$  and variance  $\Sigma_i = [\Sigma_{i,1}, \Sigma_{i,2}]^T$  for the  $i$ -th Gaussian for a certain tissue, and the probability of  $X$  belongs to this tissue can be defined as  $p_i(X) = \frac{1}{\sqrt{2\pi}|\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2}(X - \mu_i)^T \Sigma_i^{-1}(X - \mu_i)\right)$ .

The preference factors are taken into account by defining  $\omega = \begin{bmatrix} \omega_{T_1} & 0 \\ 0 & \omega_{T_2} \end{bmatrix}$ , and incorporating it in the probability, as follows:

$$p_i(X) = \frac{1}{\sqrt{2\pi}|\Sigma_i|} \exp\left(-\frac{1}{2}(X - \mu_i)^T \omega^T \Sigma_i^{-1} \omega (X - \mu_i)\right) \quad (7)$$

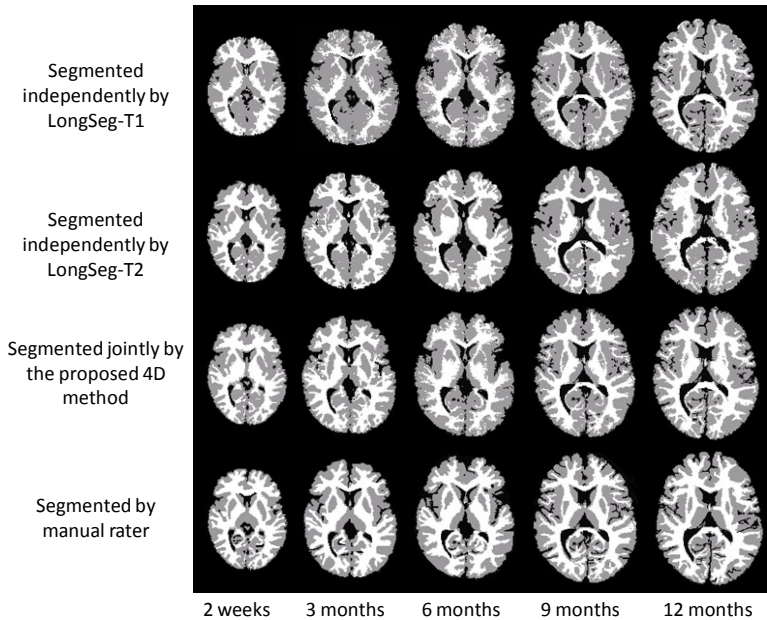
Initial means and variances are estimated based on a set of previously segmented images. Using Bayes rule, posterior probability maps can be obtained for image  $I$ , giving the probabilities of each voxel of belonging to gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). We apply the majority rule and assign the label with the maximal probability to each voxel. This newly updated segmentation label can be used to better estimate the parameters and perform another iteration of segmentation for better accuracy.

### 3 Experimental Results

To validate our proposed 4D joint registration and segmentation method, we apply it to a group of infants, scanned at 2 weeks, 3 months, 6 months, 9 months, and 12 months, with variations of same time-point scans less than 1 week. These healthy subjects were not sedated for MRI. T1 images were acquired using a 3T head-only MR scanner, with 144 sagittal slices at resolution of  $1 \times 1 \times 1 \text{ mm}^3$ . Meanwhile, T2 images of 64 axial slices were obtained at resolution of  $1.25 \times 1.25 \times 1.95 \text{ mm}^3$ . Standard preprocessing steps such as skull stripping and bias correction are performed. T2 images are affine aligned to their T1 counterparts and are resampled with a  $1 \times 1 \times 1 \text{ mm}^3$  resolution before further processing. We evaluate the segmentation accuracy on all time points by visual inspection and quantitative measurement.

For evaluation, we choose 3 subjects, and manually segment all their 5 time-points by a trained rater. The proposed segmentations are compared with manual ground truth using the Dice ratio,  $DR(L, M) = 2|L \cap M| / (|L| + |M|)$ , where  $L$  is the estimated segmentation label and  $M$  is the manual segmentation label. Dice ratio ranges from 0 to 1. For comparison, the initial segmentations are used, which are performed by a longitudinal segmentation method described in [4]. Briefly, the fuzzy segmentation result of the 12-month T1 image is used as a prior to guide the

segmentations of all other time-points, respectively. Unlike our method, it is a one-way prior propagation and only two time-points can be involved. Their results are referred to as “LongSeg-T1” and “LongSeg-T2”. Segmentations of a representative infant are shown in Fig. 4. Visually, our results show better anatomical details and are more consistent compared with the two control methods without temporal consideration, by taking manual segmentations as reference. The quantitative results are shown in Fig. 5. The accuracy of the independent segmentation on T1 images is initially low, but is eventually higher. A reverse pattern can be observed for T2 images. The proposed method is robust and achieves the best segmentation accuracy.



**Fig. 4.** Segmentation results of LongSeg-T1, LongSeg-T2, the proposed method, and a manual rater. The original T1 and T2 images are shown in Fig. 1.



**Fig. 5.** Mean Dice ratios (y-axis) of 3 subjects for the LongSeg-T1, LongSeg-T2 and the proposed method, compared with the manual segmentations. X-axis is the postnatal month when image acquired.

## 4 Conclusion

A 4D joint registration and segmentation framework is proposed in this paper. The goal is to improve segmentation of infant images from the isointense stage, and at the same time preserve temporal consistency. Spatial-temporal constrain is introduced in the construction of prior probabilistic maps, by making optimal use of both T1 and T2 images, and by adaptively borrowing knowledge from the neighboring time-points. 4D registration and multi-channel segmentation are iterated as an EM process, assisting each other to progressively generate better results. Experimental results indicate that our method yields the highest agreement with the manual rater, and outperforms other methods in comparison.

Different with the longitudinal segmentation method in [4], the proposed method considers significantly more time-points, with the ability to take advantage of information from both temporally increasing and decreasing directions. It also incorporates a more principled structural-temporal constraint, effectively gathering more information spatially and temporally for accurate segmentation of not only neonatal images but also images at the 6 months time-point, which typically have images with very low tissue contrast. The proposed method will be further tested using a large set of longitudinal infant images currently acquired in the authors' institute, and is expected to facilitate longitudinal infant brain development research by providing an effective means for segmenting infant images from the isointense stage with temporal consistency.

## References

1. Knickmeyer, R.C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J.K., Hamer, R.M., Lin, W., Gerig, G., Gilmore, J.H.: A structural MRI study of human brain development from birth to 2 years. *J. Neurosci.* 28, 12176–12182 (2008)
2. Prastawa, M., Gilmore, J.H., Lin, W., Gerig, G.: Automatic segmentation of MR images of the developing newborn brain. *Medical Image Analysis* 9, 457–466 (2005)
3. Xue, H., Srinivasan, L., Jiang, S., Rutherford, M., Edwards, A.D., Rueckert, D., Hajnal, J.V.: Automatic segmentation and reconstruction of the cortex from neonatal MRI. *Neuroimage* 38, 461–477 (2007)
4. Shi, F., Fan, Y., Tang, S., Gilmore, J.H., Lin, W., Shen, D.: Neonatal brain image segmentation in longitudinal MRI studies. *Neuroimage* 49, 391–400 (2010)
5. Dietrich, R.B., Bradley, W.G., Zaragoza, E.J.t., Otto, R.J., Taira, R.K., Wilson, G.H., Kangaroo, H.: MR evaluation of early myelination patterns in normal and developmentally delayed infants. *American Journal of Roentgenology* 150, 889–896 (1988)
6. Shattuck, D.W., Leahy, R.M.: Automated graph-based analysis and correction of cortical volume topology. *IEEE Transactions on Medical Imaging* 20, 1167–1177 (2001)
7. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging* 17, 87–97 (1998)
8. Ashburner, J., Friston, K.J.: Unified segmentation. *Neuroimage* 26, 839–851 (2005)
9. Shen, D., Davatzikos, C.: Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping. *Neuroimage* 21, 1508–1517 (2004)
10. Sundar, H., Litt, H., Shen, D.: Estimating myocardial motion by 4D image warping. *Pattern Recognition* 42, 2514–2526 (2009)