

# Multi-atlas and Multi-modal Hippocampus Segmentation for Infant MR Brain Images by Propagating Anatomical Labels on Hypergraph

Pei Dong, Yanrong Guo, Dinggang Shen, and Guorong Wu<sup>(✉)</sup>

Department of Radiology and BRIC,  
University of North Carolina at Chapel Hill,  
Chapel Hill, NC, USA

{pei.dong, yrguo, dgshen, guorong\_wu}@med.unc.edu

**Abstract.** Accurate segmentation of hippocampus from infant magnetic resonance (MR) images is very important in the study of early brain development and neurological disorder. Recently, multi-atlas patch-based label fusion methods have shown a great success in segmenting anatomical structures from medical images. However, the dramatic appearance change from birth to 1-year-old and the poor image contrast make the existing label fusion methods less competitive to handle infant brain images. To alleviate these difficulties, we propose a novel multi-atlas and multi-modal label fusion method, which can unanimously label for all voxels by propagating the anatomical labels on a hypergraph. Specifically, we consider not only all voxels within the target image but also voxels across the atlas images as the vertexes in the hypergraph. Each hyperedge encodes a high-order correlation, among a set of vertexes, in different perspectives which incorporate (1) feature affinity within the multi-modal feature space, (2) spatial coherence within target image, and (3) population heuristics from multiple atlases. In addition, our label fusion method further allows those reliable voxels to supervise the label estimation on other difficult-to-label voxels, based on the established hyperedges, until all the target image voxels reach the unanimous labeling result. We evaluate our proposed label fusion method in segmenting hippocampus from T1 and T2 weighted MR images acquired from at 2-week-old, 3-month-old, 6-month-old, 9-month-old, and 12-month-old. Our segmentation results achieves improvement of labeling accuracy over the conventional state-of-the-art label fusion methods, which shows a great potential to facilitate the early infant brain studies.

## 1 Introduction

The human brain undergoes a rapid physical growth and fast functional development during the first year of life. In order to characterize such dynamic changes *in vivo*, accurate segmentation of anatomical structures from the MR images is very important

---

Dong and Guo-These authors contributed equally to this work.

in imaging-based brain development studies. Since hippocampus plays an important role in learning and memory function, many studies aim to find the imaging markers around hippocampus. Unfortunately, the poor image contrast and dramatic appearance change make the segmentation of hippocampus from the infant images in the first year of life very challenging, due to the dynamic white matter myelination progress [1]. Since different modalities convey diverse imaging characteristics at different brain development phases, integration of multi-modal imaging information can significantly improve the segmentation accuracy [2]. However, the modulation of each modality is hard-coded and subjective to the expert’s experience.

In medical image analysis area, there has been a recent spike in multi-atlas patch based segmentation methods [3–5], which first register all atlas images to the underlying target image and then propagate the labels from the atlas domain to the target image. The assumption behind is that two voxels should bear the same anatomical label if their local appearances are similar. However, a critical issue in current multi-atlas patch based label fusion methods is that the labels are determined separately at each voxel. As a result, there is no guarantee that labeled anatomical structures are spatially consistent. On the other hand, conventional graph-cut based methods [6] can jointly segment the entire ROI by finding a minimum graph cut. However, only the image information within the target image is utilized during this segmentation.

To combine the power of multi-atlas and graph-based approaches, we propose a novel multi-atlas and multi-modal label fusion method for segmenting the hippocampus from the infant brain images in the first year of life. Hypergraph has shown its superiority in image retrieval [7, 8] and recognition [9]. Our idea is to use hypergraph to leverage (1) the information integration from multiple atlases and multiple imaging modalities, and (2) the label fusion for all target image voxels under consideration.

Specifically, we regard all voxels within the target image and the atlas images as the vertexes in the graph. In hypergraph, we generalize the concept of conventional graph edge (only connects two vertexes at a time) to the hyperedge (groups a set of vertexes simultaneously), in order to reveal the high-order correlations for more than two voxels. In general, our hyperedges in the multi-atlas scenario encode three types of voxels correlations: (1) *feature affinity*: only the vertexes with similar appearance are connected by hyperedge, (2) *spatial coherence*: the vertexes located in a certain neighborhood from target image belong to the same hyperedge, and (3) *atlas correspondence*: each vertex from target image and its corresponding vertexes across all atlas images form the hyperedge. It is worth noting that the hypergraph is very flexible to incorporate the multi-modal information by constructing the above three types of hyperedges w.r.t. different imaging modalities. After constructing the hypergraph, the vertexes from both the atlas image voxels and the target image voxels with high labeling confidence are considered bearing the known labels. Thus, our label fusion method falls into the semi-supervised hypergraph learning framework, i.e., the latent labels on the remaining voxels in the target image are influenced by the connected vertexes with known labels. The principle to propagate the labels is that the vertexes sitting in the same hyperedge should have the same label, with the minimal discrepancy of labels on the vertexes with existing labels before and after label propagation.

Our proposed method has been comprehensively evaluated on segmenting hippocampus from T1 and T2 weighted MR images at 2-week-old, 3-month-old, 6-month-old, 9-month-old and 12-month-old. The segmentation result shows a great improvement compared to the state-of-the-art method [3–5] in terms of labeling accuracy.

## 2 Method

Given the atlases (including T1, T2 weighted MR images and the ground truth label of hippocampus), the goal is to segment the hippocampus from the unlabeled target image. Since the T1 and T2 weighted images are acquired from the same subject at the same time, it is not difficult to register them to the same space. Due to the dynamic appearance change and poor image resolution, it is challenging to apply deformable image registration for the infant images in the first year of life. Thus, only linear registration is used to map each atlas to the underlying target image domain. After that, our label fusion method consists of two main steps: (1) hypergraph construction (Sect. 2.1) and (2) label propagation (Sect. 2.2), as detailed below.

### 2.1 Hypergraph Construction

We assume that  $N$  atlas images, each with T1 and T2 weighted MR images, are used to label the target image. The hypergraph is constructed to accommodate the complete information from both the multi-atlas and multi-modalities images, where the hypergraph is denoted as  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$  with the vertex set  $\mathcal{V}$ , the hyperedge set  $\mathcal{E}$  and its weight  $\mathbf{W}$ . In order for computation efficiency, only the voxels within a bounding box (shown in the left column of Fig. 1 which covers the hippocampus with enough margin) are used as the vertexes, instead of using the voxels from the entire image.

**Hypergraph Initialization:** Assume the number of voxels inside the bounding box is  $q$ . The hypergraph vertexes are built on every voxel from the bounding boxes of the target image and the atlas images. For each atlas image, we vectorize the voxels in the bounding box into a column vector  $\mathcal{V}_s$  with the length of  $q$ , where  $s = 1, \dots, N$ . With the same vectorising order, we can obtain the column vector of graph vertexes from target image, denoted by  $\mathcal{U}$ . Thus, the vertex set  $\mathcal{V}$  is eventually the combination of voxels from the target image and all atlas images, i.e.,  $\mathcal{V} = \{\mathcal{U}, \mathcal{V}_1, \dots, \mathcal{V}_N\}$ . It is worth noting that each element in  $\mathcal{V}_s$  bears the known label and the label on each element of  $\mathcal{U}$  is currently unknown.

It is clear that the voxels along the interface of multiple structures are more difficult to determine the labels than other points. In light of this, we go one step further to classify the elements in  $\mathcal{U}$  into two categories based on their difficulties in labeling. To achieve it, we use an existing label fusion method, such as the majority voting, to predict the label of each element in  $\mathcal{U}$  as well as the confidence value in terms of the voting predominance. If the influence for voting one label dominates the other labels, we regard the labeling result on the underlying target image voxel has high confidence

to determine its label. Otherwise, we regard the underlying target image voxel needs other heuristics to find its label. Thus, we can divide  $\mathcal{U}$  into two parts,  $\mathcal{U}^L$  (low confidence) and  $\mathcal{U}^H$  (high confidence), where the number of elements in  $\mathcal{U}^L$  and  $\mathcal{U}^H$  is  $q^L$  and  $q^H$  ( $q = q^L + q^H$ ), respectively.

The advantage of separate  $\mathcal{U}$  to  $\mathcal{U}^L$  and  $\mathcal{U}^H$  is that it allows the label propagation from not only the atlas images but also the some reliable regions of the target image, which are more specific to label fusion of underlying image. Specifically, for each element in  $\mathcal{U}^H$  and  $\{\mathcal{V}_s | s = 1, \dots, N\}$ , we assign the label with either 0 (background) or 1 (hippocampus). For each element in  $\mathcal{U}^L$ , the label value is assigned to 0.5 since the label is uncertain. Following the same order of  $\mathcal{V}$ , we stack all the label values into a label vector  $\mathbf{y}$ . As shown in the left bracket in Fig. 1, the first part of  $\mathbf{y}$  is the uncertain labels for voxels  $\mathcal{U}^L$  (in gray), followed by the known labels from  $\mathcal{U}^H$  and  $\{\mathcal{V}_s | s = 1, \dots, N\}$  (hippocampus in white and background in black). Thus, our hypergraph based label fusion turns into a semi-supervised learning scenario, i.e., optimize for a new label vector  $\mathbf{f}$  (in the right of Fig. 1) which should be (1) as close to  $\mathbf{y}$  as possible and (2) propagate the known labels from  $\mathcal{U}^H$  and  $\{\mathcal{V}_s | s = 1, \dots, N\}$  to the difficult-to-label vertexes  $\mathcal{U}^L$ . The leverage of label propagation is a set of hyperedges, as detailed next.

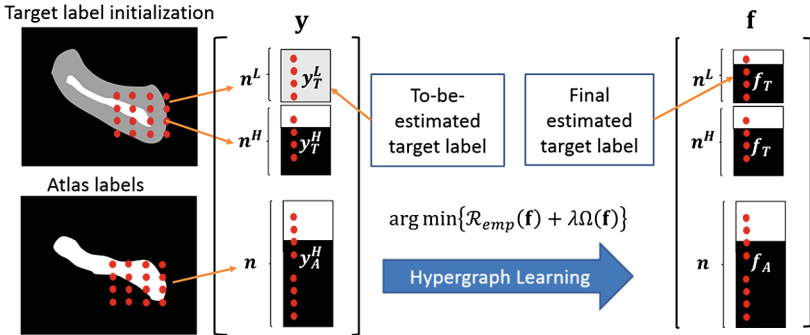


Fig. 1. The framework of hypergraph learning based label propagation.

**Hyperedge Construction:** Since the goal is to estimate the latent labels for the difficult-to-label voxels in the target image, we only construct hyperedges centered at each vertex  $u \in \mathcal{U}^L$  in the low confidence voxel set. In total, we constructed  $2 \times 3$  types of hyperedges w.r.t. two imaging modalities and three measurements. Specifically, for each modality, we construct the following three hyperedges on each vertex  $u$ :

Feature Affinity (FA) hyperedge  $e^{FA}$ . It consists of  $K$  nearest vertexes ( $v \in \mathcal{V}$ ), where their patchwise similarities w.r.t. the underlying vertex  $u$  are the highest  $K$  hits in the feature space.

Local Coherence (LC) hyperedge  $e^{LC}$ . It consists of all the vertexes  $v \in \mathcal{U}$  from the target image and being located within a spatial neighborhood of  $u$ .

Atlas Correspondence (AC) hyperedge  $e^{AC}$  It consists of all vertexes  $v \in \{\mathcal{V}_s | s = 1, \dots, N\}$  across the atlas images at the corresponding spatial locations w.r.t.  $u$ .

Since each of the above hyperedge is centered by  $u$ , we call  $u$  as the owner of hyperedge  $e$  (we omit the superscript of  $e$  hereafter).

**Incidence Matrix Construction:** After the construction of hyperedges for each  $u$ , we build a incidence matrix  $\mathbf{H}$  with row representing the vertexes and column representing the hyperedges  $e$ , which encode all the information within the hypergraph  $\mathcal{G}$ . Each entry  $\mathbf{H}(v, e)$  in  $\mathbf{H}$  measures the affinity between the each vertex  $v$  and the owner  $u$  of the hyperedge  $e \in \mathcal{E}$ :

$$\mathbf{H}(v, e) = \begin{cases} \exp\left(-\frac{\|\mathbf{p}(v) - \mathbf{p}(u)\|_2^2}{\sigma^2}\right) & \text{if } v \in e \\ 0 & \text{if } v \notin e \end{cases} \quad (1)$$

where the  $\|\cdot\|_2$  is the  $L_2$  norm computed between intensity image patch  $\mathbf{p}(v)$  and  $\mathbf{p}(u)$  for vertex  $v$  and the hyperedge owner  $u$ .  $\sigma$  is the averaged patchwise similarity between  $u$  and all vertexes within the hyperedge  $e$ . For simplicity, each hyperedge is initialized with an equal weight,  $\mathbf{W}(e) = 1$ . The degree of a vertex  $v$  is defined as  $d(v) = \sum_{e \in \mathcal{E}} \mathbf{w}(e) \mathbf{H}(v, e)$ , and the degree of hyperedge is defined as  $\delta(e) = \sum_{v \in \mathcal{V}} \mathbf{H}(v, e)$ . Thus, two diagonal matrices  $\mathbf{D}_v$  and  $\mathbf{D}_e$  can be formed with each entry along the diagonal using the vertex degree and hyperedge degree, respectively.

According to this formulation, the construction of the hypergraph captures the high-order relationship among the all the vertices across different image modalities, feature affinities, local neighborhood coherences and the atlas correspondences. And any two vertexes within the same hyperedge should have similar labels. Following gives the detailed hypergraph learning for label propagation.

## 2.2 Labels Propagation via Hypergraph Learning

Given the initialization on the hypergraph, we employed a semi-supervised learning method to perform the label fusion on the constructed hypergraph. The objective function [10] is defined as:

$$\arg \min_{\mathbf{f}} \{ \mathcal{R}_{emp}(\mathbf{f}) + \lambda \Omega(\mathbf{f}) \} \quad (2)$$

where  $\mathbf{f}$  is the likelihood of hippocampus on each vertex  $v \in \mathcal{V}$ . The first term  $\mathcal{R}_{emp}(\mathbf{f})$  is an empirical loss, which prevent the dramatic change for those the high probability values in  $\mathbf{f}$ . The second term  $\Omega(\mathbf{f})$  is a regularization term on the hypergraph, which restrict the labels at similar range for those vertexes within a same hyperedge. The  $\lambda$  is a positive weighing parameter between the two terms.

The first empirical loss term  $\mathcal{R}_{emp}(\mathbf{f})$  is defined by

$$\mathcal{R}_{emp}(\mathbf{f}) = \|\mathbf{f} - \mathbf{y}\|^2 \quad (3)$$

The empirical loss term is designed for the minimization of the differences before and after label fusion.

The second term  $\Omega(\mathbf{f})$  is a regularizer on the hypergraph, which defined as follows:

$$\Omega(\mathbf{f}) = \frac{1}{2} \sum_{e \in \mathcal{E}} \sum_{u, v \in \mathcal{V}} \frac{\mathbf{W}(e) \mathbf{H}(u, e) \mathbf{H}(v, e)}{\delta(e)} \left( \frac{\mathbf{f}(u)}{\sqrt{d(u)}} - \frac{\mathbf{f}(v)}{\sqrt{d(v)}} \right)^2 = \mathbf{f}^T (\mathbf{I} - \Theta) \mathbf{f} = \mathbf{f}^T \Delta \mathbf{f} \quad (4)$$

where  $\Theta = \mathbf{D}_v^{-\frac{1}{2}} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-\frac{1}{2}}$ , and  $\Delta$  can be viewed as normalized hypergraph Laplacian matrix. Here, the regulation term  $\Omega(\mathbf{f})$  constrains the labels of vertexes in the same hyperedge should held a similar value during the label propagation. Thus, the objective function, Eq. (2), can be rewritten as:

$$\arg \min_{\mathbf{f}} \left\{ \|\mathbf{f} - \mathbf{y}\|^2 + \lambda \mathbf{f}^T \Delta \mathbf{f} \right\} \quad (5)$$

By differentiating the objective function with respect to  $\mathbf{f}$ , the optimal  $\mathbf{f}$  can be computed iteratively as below:

$$\mathbf{f} = \left( \mathbf{I} + \frac{1}{\lambda} (\mathbf{I} - \Theta) \right)^{-1} \mathbf{y} \quad (6)$$

Given the estimated  $\mathbf{f}$ , the anatomical label on each difficult-to-label target voxel  $u \in \mathcal{U}^L$  can be determined by

$$\begin{cases} \text{Hippocampus} & f(u) > 0.5 \\ \text{Background} & \text{otherwise} \end{cases} \quad (7)$$

It should be noted that using the semi-supervised learning to perform the binary classification on the hypergraph allows prediction of the entire target image voxels simultaneously, while combining several correlations across all the voxels between the target image and atlas images.

## 3 Experiments

### 3.1 Data Acquisition and Preprocessing

In the experiments, MR images of 10 healthy infant subjects are acquired from a Siemens head-only 3T scanner. For each subject, both T1- and T2-weighted MR images were acquired in five data sets at 2 weeks, 3 months, 6 months, 9 months and 12 months of age. T1-weighted MR images were acquired with 144 sagittal slices at a

resolution of  $1 \times 1 \times 1 \text{ mm}^3$ , while T2-weighted MR images were acquired with 64 axis slices at resolution of  $1.25 \times 1.25 \times 1.95 \text{ mm}^3$ . For each subject, the T2-weighted MR image is linearly aligned to the T1-weighted MR image at the same age and then further resampled to  $1 \times 1 \times 1 \text{ mm}^3$ . Standard preprocessing was performed including skull stripping [11], intensity inhomogeneity correction [12]. The manual segmentations of the hippocampal regions for all 10 subjects are available and used as ground-truth for evaluation.

### 3.2 Evaluation of the Proposed Method

For all the experiments, the patch size is set as  $5 \times 5 \times 5$  voxels. The number of nearest neighborhood vertexes  $K$  is 10 in constructing  $e^{FA}$ . The spatial neighborhood in constructing  $e^{LC}$  and  $e^{AC}$  is set to  $3 \times 3 \times 3$  voxel. Parameter  $\lambda$  in Eq. (2) is 10.

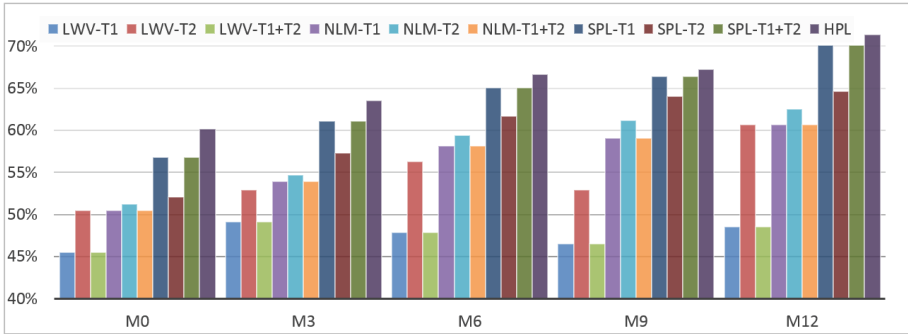
To evaluate the performance of the proposed method, we adopted the leave-one-out cross-validation. In each cross-validation step, one subject is used as target images and the remaining 9 subjects were used as the atlas images. Our proposed hypergraph patch labeling (HPL) method is compared with three state-of-the-art multi-atlas patch-labeling methods: local-weighted majority voting (LMV) [3], non-local mean (NLM) [4] and sparse patch labeling (SPL) [5].

Table 1 gives the average dice ratio of four comparison methods for 2-week, 3-month, 6-month, 9-month and 12-month data respectively. It can be seen that our method can achieve better segmentation accuracy than all other three counterpart methods. Figure 2 further shows the impact of using our proposed method for segmenting the infant hippocampus compared with other different labeling methods under different modalities.

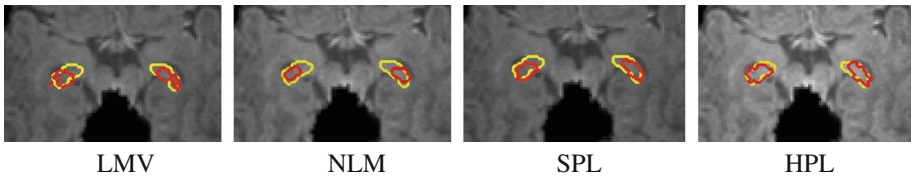
Finally, Fig. 3 shows the final segmentation results by the four methods and their corresponding manually segmentations for one typical subject at 2-week-old. Through visual inspection, our estimated labeling results (bottom line) are closer to the ground truth (yellow contours).

**Table 1.** The average dice ratio and standard deviation of four comparison methods: local-weighted majority voting, non-local mean, sparse patch labeling and hypergraph patch labeling for 2 week, 3 month, 6 month, 9 month and 12 month data. (\* indicates the significant improvement of HPL over other compared methods ( $p < 0.05$ ), T1 + T2).

	LMV	NLM	SPL	HPL
2 week	0.45 $\pm$ 0.12*	0.51 $\pm$ 0.12	0.57 $\pm$ 0.10	0.60 $\pm$ 0.11
3 month	0.49 $\pm$ 0.09*	0.54 $\pm$ 0.10*	0.61 $\pm$ 0.04	0.64 $\pm$ 0.04
6 month	0.48 $\pm$ 0.09*	0.58 $\pm$ 0.11	0.65 $\pm$ 0.06	0.67 $\pm$ 0.07
9 month	0.46 $\pm$ 0.07*	0.59 $\pm$ 0.08*	0.66 $\pm$ 0.04	0.67 $\pm$ 0.04
12 month	0.48 $\pm$ 0.10*	0.61 $\pm$ 0.05*	0.70 $\pm$ 0.04	0.71 $\pm$ 0.04



**Fig. 2.** Comparison of the average dice ratio of the ten infant subjects between different labeling methods across five time points: 2-week (M0), 3-month (M3), 6-month (M6), 9-month (M9) and 12-month (M12)



**Fig. 3.** Infant hippocampus segmentation comparison between the automatic segmentation using the four methods (red circles) and the ground truth (yellow circles) (Color figure online)

## 4 Discussion and Conclusion

With our proposed method, we achieved a better segmentation result with fewer atlases compared with the state-of-the-art patch-based method. The reasons are due to the followings: (1) the label propagation based on the hypergraph can adaptively combine the complementary information both from T1- and T2- weighed images; (2) for each image modality, three types of voxels correlations are considered, i.e. feature affinity, local spatial neighborhood and atlas corresponding spatial regions, which extensively exploits the high order correlations among a group of voxels; (3) Instead of labeling each voxel independently as in the existing methods, our semi-supervised hypergraph learning method performs a global optimization to predict the anatomical labels for all the voxels in the target image simultaneously. (4) Through the label propagation process, we allow the high confidence voxels to guide the nearby low confidence voxels, which provides valuable heuristics to overcome the uncertainty in label fusion.

In this paper, we propose a multi-atlas and multi-modal label fusion method for the segmentation of the hippocampus from infant MR images. We combine the advantages of conventional multi-atlas and graph-based approaches by encoding the feature affinity within the multi-modal feature space, spatial coherence, and the atlas information via hypergraph. Then, the whole label fusion procedure falls into the semi-supervised hypergraph learning framework, where the estimation of the latent anatomical labels is



adaptively influenced by the connected counterparts in the hypergraph. The experiment results show more accurate hippocampus labeling results from MR images in the different phases of first year of life, using our proposed method with comparison to the state-of-the-art methods.

## 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. Off. J. Soc. 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. Isgum, I., Staring, M., Rutten, A., Prokop, M., Viergever, M.A., van Ginneken, B.: Multi-atlas-based segmentation with local decision fusion—application to cardiac and aortic segmentation in CT scans. *IEEE Trans. Med. Imaging* **28**, 1000–1010 (2009)
4. Coupé, P., Manjón, J.V., Fonov, V., Pruessner, J., Robles, M., Collins, D.L.: Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage* **54**, 940–954 (2011)
5. Zhang, D., Guo, Q., Wu, G., Shen, D.: Sparse patch-based label fusion for multi-atlas segmentation. In: Yap, P.-T., Liu, T., Shen, D., Westin, C.-F., Shen, L. (eds.) *MBIA 2012*. LNCS, vol. 7509, pp. 94–102. Springer, Heidelberg (2012)
6. Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888–905 (2000)
7. Gao, Y., Ji, R., Cui, P., Dai, Q., Hua, G.: Hyperspectral image classification through bilayer graph-based learning. *IEEE Trans. Image Process.* **23**(7), 2769–2778 (2014)
8. Gao, Y., Wang, M., Zha, Z.-J., Shen, J., Li, X., Wu, X.: Visual-textual joint relevance learning for tag-based social image search. *IEEE Trans. Image Process.* **22**(1), 363–376 (2013)
9. Gao, Y., Wang, M., Tao, D., Ji, R., Dai, Q.: 3-D object retrieval and recognition with hypergraph analysis. *IEEE Trans. Image Process.* **21**(9), 4290–4303 (2012)
10. Zhou, D., Huang, J., Schölkopf, B.: Learning with hypergraphs: clustering, classification, and embedding. In: *Proceedings of NIPS*, vol. 19, pp. 1601–1608 (2006)
11. Shi, F., Wang, L., Dai, Y., Gilmore, J.H., Lin, W., Shen, D.: LABEL: pediatric brain extraction using learning-based meta-algorithm. *NeuroImage* **62**, 1975–1986 (2012)
12. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imaging* **17**, 87–97 (1998)