

HHS Public Access

Author manuscript

Patch Based Tech Med Imaging (2016). Author manuscript; available in PMC 2018 October 05.

Published in final edited form as:

Patch Based Tech Med Imaging (2016). 2016; 9993: 34-42. doi:10.1007/978-3-319-47118-1_5.

Consistent Multi-Atlas Hippocampus Segmentation for Longitudinal MR Brain Images with Temporal Sparse Representation

Lin Wang^{1,2}, Yanrong Guo², Xiaohuan Cao^{2,3}, Guorong Wu², and Dinggang Shen²

¹School of Information Science and Technology, Northwest University, Xi'an, China

²Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

³School of Automation, Northwestern Polytechnical University, Xi'an, China

Abstract

In this paper, we propose a novel multi-atlas based longitudinal label fusion method with temporal sparse representation technique to segment hippocampi at all time points simultaneously. First, we use groupwise longitudinal registration to simultaneously (1) estimate a group-mean image of a subject image sequence and (2) register its all time-point images to the estimated group-mean image consistently over time. Then, by registering all atlases with the group-mean image, we can align all atlases longitudinally consistently to each time point of the subject image sequence. Finally, we propose a longitudinal label fusion method to propagate all atlas labels to the subject image sequence by simultaneously labeling a set of temporally-corresponded voxels with a temporal consistency constraint on sparse representation. Experimental results demonstrate that our proposed method can achieve more accurate and consistent hippocampus segmentation than the state-of-the-art counterpart methods.

1 Introduction

The hippocampus plays a crucial role in memory and spatial navigation function of brain [1]. The structural change of hippocampus over time is highly related to many neurodegenerative diseases, such as Alzheimer's disease (AD). As a characteristic feature of AD, hippocampal atrophy is considered as a potential biomarker for the diagnosis and assessment of AD in magnetic resonance (MR) imaging based neuro-science studies [2, 3]. In order to measure the hippocampal atrophy over time, accurate quantization of hippocampal volumes from serial structural three-dimensional (3D) MR images is required. To this end, it is important to accurately and temporally consistently segment fourdimensional (3D+t) hippocampus from longitudinal structural MR images.

Many automatic segmentation methods have been proposed to segment 3D hippocampus independently from MR images of different time points [4–8]. In this case, hippocampus

segmentation from longitudinal MR images is decomposed into a series of separate hippocampus segmentation from each 3D MR image. However, due to various reasons such as variant noises and hippocampal tissue contrast in the acquired longitudinal MR images, 3D segmentation methods, when applied to longitudinal MR images, have limited temporal consistency for segmented longitudinal hippocampi. To improve segmentation consistency from longitudinal MR images, several longitudinal segmentation methods have been proposed [9–12]. Wolz et al. [9] proposed a 4D graph-cut based method to simultaneously segment longitudinal MR images. By using a 4D graph to represent longitudinal MR data, the method segmented MR images at all time points by solving the min-cut/max-flow problem on the 4D graph. Chincarini et al. [12] presented a hippocampal segmentation method by integrating longitudinal information. They implemented longitudinal analysis with four progressive steps, and addressed the impact of these steps on longitudinal performance of hippocampal volume measurements for early detection of AD. However, due to large variance of noise level and intensity bias field across different time points in the longitudinal MR images, consistent hippocampus segmentation of serial MR images remains a challenging problem.

Accordingly, in this paper, we propose a 3D+t hippocampus segmentation method for longitudinal MR brain images, by integrating temporal sparse representation within the multi-atlas patch-based label fusion framework. First, we use the groupwise longitudinal image registration toolbox (GLIRT) [13] to simultaneously (1) estimate a subject-specific group-mean image and (2) register all time-point images of subject image sequence consistently to the estimated group-mean image. Then, by registering all atlas images to the estimated group-mean image, we can align all atlases longitudinally consistently to each time point of the subject image sequence. Thus, given the temporal correspondence in subject image sequence, we can form a 3D+t image patch at each location of subject brain. Then, we can use temporal sparse representation technique to simultaneously determine labels for all time points of the subject image sequence by propagating labels from all aligned atlas images. Experimental results on both simulated and real longitudinal MR images demonstrate that our proposed method can achieve more accurate and consistent hippocampus segmentation than the state-of-the-art multi-atlas label fusion methods, which often apply label fusion for each time point independently.

2 Method

Before describing our method in detail, we first introduce some mathematical descriptions. First, a longitudinal subject image sequence (also namely 3D+t subject image) is denoted as $\{T'_t\}_{t=1}^N$, where T'_t is a 3D image at time point $t(t \in \{1, ..., N\})$. Then, *M* 3D atlas images are denoted as $I'^{(1)}, ..., I'^{(M)}$ with their corresponding hippocampal label images denoted as $L'^{(1)}, ..., L'^{(M)}$. So, our goal is to segment 3D+t subject image, i.e., to automatically estimate a hippocampal label image sequence $\{L'_t\}_{t=1}^N$ corresponding to the 3D+t subject image $\{T'_t\}_{t=1}^N$ as illustrated in Fig. 1, where L'_t denotes a 3D subject label image at time point $t(t \in \{1, ..., N\})$.

2.1 Temporal Sparse Representation

Unbiased Groupwise Registration—The first step in our method is to estimate temporal correspondences along different time points of the subject image sequence. Here, we use GLIRT [13] to achieve this goal. Unlike other pairwise registration methods, which need choose a reference template, the groupwise registration is free of template selection, and can thus build an unbiased subject-specific group-mean image and simultaneously register all time-point images to this group-mean image. Assuming $\{\varphi'_t\}_{t=1}^N$ as the deformation fields for *N* time points, the original 3D+t subject image $\{T'_t\}_{t=1}^N$ can be transformed to the group-mean image space as $\{T'_t\}_{t=1}^N$, which is illustrated in Fig. 1. Here, $T_t = \varphi'_t(T'_t)$..

Furthermore, all 3D atlas images are registered to the estimated group-mean image, first by affine registration [14] with 12 degrees of freedom and then by deformable registration [15]. By following the same estimated *affine transformation matrix* and *deformation field*, the corresponding 3D hippocampal label image of each atlas can be also registered to the estimated group-mean image space. As shown in Fig. 1, $I^{(1)}$, ..., $I^{(M)}$ and $L^{(1)}$, ..., $L^{(M)}$ denote the *M* aligned 3D atlas images and their corresponding aligned 3D label images in the estimated group-mean image space, respectively.

Temporal Sparse Patch-Based Representation—For each voxel x in the group-mean image, its 3D+t subject patch can be extracted as $\{\alpha_{(x,t)}\}_{t=1}^{N}$, where $\alpha_{(x,t)}$ represents a 3D patch centered at voxel x of the aligned 3D+t subject image at time point t. Let n(x) denote a spatial neighborhood of x. All candidate atlas patches within the search neighborhood n(x) of the aligned atlas images $\{I^{(m)}\}_{m=1}^{M}$ are denoted as $\{\beta_{z}^{(m)}|z \in n(x), m=1, ..., M\}$, along with their corresponding center voxel labels $\{I_{z}^{(m)}|z \in n(x), m=1, ..., M\}$. The total number of atlas patches $\{\beta_{z}^{(m)}|z \in n(x), m=1, ..., M\}$ used to label the 3D+t subject patch $\{\alpha_{(x,t)}\}_{t=1}^{N}$ is $Q = M \times |n(x)|$, where |n(x)| denotes the cardinality of n(x).

After rearranging $\beta_z^{(m)}$ and $a_{(x,t)}$ into a *d*-dimensional column vector $b_z^{(m)}$ and $a_{(x,t)}$, respectively, where *d* is the number of voxels in each 3D patch, our temporal sparse representation can be formulated as a problem of finding optimal sparse representation for the 3D+t subject patch vector $\{a_{(x,t)}\}_{t=1}^{N}$ by using all atlas patch vectors $\{b_z^{(m)}\}_{m=1}^{M}$ as follows [6]:

$$\{ \mathbf{w}_{(\mathbf{x},t)} \}_{t=1}^{N} \left\| + \lambda_2 \sum_{t=1}^{N-1} \left\| \mathbf{w}_{(\mathbf{x},t)} - \mathbf{w}_{(\mathbf{x},t+1)} \right\|_1 \right\} s \cdot t \cdot w_{(\mathbf{x},t)} \ge 0, \, \forall t = 1, ..., N$$

where $B_x \in \mathbb{R}^{d \times Q}$ denotes a dictionary matrix constructed by arranging $\{b_z^{(m)}\}_{m=1}^M$ column by column, $w_{(x,t)} \in \mathbb{R}^Q$ denotes a weight vector by arranging all non-negative weights $\{w_{(x,t),z}^{(m)} | w_{(x,t),z}^{(m)} \ge 0, z \in n(x), m = 1, ..., M\}$ into a column vector, and $w_{(x,t),z}^{(m)}$ is the representation coefficient of the patch vector $b_z^{(m)}$ of the *m*-th atlas image in constructing the patch vector $a_{(x,t)}$ of the subject image at time point *t*. The first term in Eq. (1) is the reconstruction discrepancy. The second term, which is equivalent to the l_1 -norm, enforces sparsity in $w_{(x,t)}$. The third term is the temporal fused smoothness term, used to constrain the temporal consistency of two successive sparse representation vectors ($w_{(x,t)}$ and $w_{(x,t+1)}$). λ_1 and λ_2 are the two weighting parameters used to balance the contributions from the second and third terms. The objective function in Eq. (1) can be solved by the fast proximal gradient method [16].

2.2 Multi-Atlas Based Label Fusion with Temporal Sparse Representation

Once the temporal sparse code $\{\hat{w}_{(x,t)}\}_{t=1}^{N}$ is estimated by solving the optimization problem in Eq. (1), the label at the voxel (x, t) of the aligned subject image T_t can be obtained by the multi-atlas based label fusion method by combining the center voxel labels $\{I_z^{(m)} | z \in n(x), m = 1, ..., M\}$ using the estimated temporal sparse code $\{\hat{w}_{(x,t)}\}_{t=1}^{N}$.

By following the same order of the dictionary matrix B_x , $\{ l_z^{(m)} | z \in n(x), m = 1, ..., M \}$ is constructed as a label vector $l_x(l_x \in \mathbb{R}^{1 \times Q})$. Supposing there are *P* possible labels $\{L_1, ..., L_p, ..., L_P\}$ in the atlases, the label at the voxel (x, t) of the aligned subject image T_t can be determined by:

$$\hat{L}_{(\mathbf{x},t)} = \arg \min_{L_p, p = 1, ..., P} \left\{ \sum_{j=1}^{Q} \widehat{w}_{(\mathbf{x},t), j} \cdot \delta(l_{\mathbf{x}, j}, L_p) \right\}, \ \forall t = 1, ..., N$$
(2)

where $\hat{w}_{(x,i),j}$ and $l_{x,j}$ are the *j*-th components of $\hat{w}_{(x,i)}$ and l_x , respectively, and the function $\delta\{l_{x,j}, L_p\}$ is equal to 1 if $l_{x,j} = L_p$ and 0 otherwise.

Patch Based Tech Med Imaging (2016). Author manuscript; available in PMC 2018 October 05.

(1)

After determining the label $\hat{L}_{(x,t)}$ at each voxel (x, t) of aligned 3D+t subject image $\{T_t\}_{t=1}^N$, the aligned 3D+t subject label image $\{L_t\}_{t=1}^N$ in the group-mean image space can be obtained. Then, we can obtain the 3D+t subject label image $\{L'_t\}_{t=1}^N$ corresponding to the input 3D+t subject image $\{T'_t\}_{t=1}^N$ by transforming $\{L_t\}_{t=1}^N$ back to the subject image space by following $L'_t = \varphi_t^{'-1}(L_t)$, where $\varphi_t^{'-1}$ is the inverse transformation field of φ_t' .

3 Experimental Results

In this section, we evaluate our proposed longitudinal (3D+t) hippocampus segmentation method on both simulated and real longitudinal MR brain image datasets. Specifically, 10 subjects with simulated atrophy in hippocampi, and 12 subjects with each subject having three MR images acquired at three time points in the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://www.adni-info.org/) are used, respectively. All the images in both datasets are the T1-weighted MR images, which are processed to have the same size and same resolution of $256 \times 256 \times 256$ and $1 \times 1 \times 1$ mm³, respectively. The proposed method is compared to two state-of-the-art label fusion methods, namely nonlocal patchbased method (Non-local) [5] and sparse patch-based method (SPBM) [6]. To be fair, a similar unbiased groupwise registration is applied to the two comparison methods. That is, subject image is registered to the estimated group-mean image of subject image sequence by GLIRT, and all atlases are registered to the estimated group-mean image sequentially by affine registration and deformable registration. The final segmentations are evaluated in each subject's own space. In the following experiments, the patch size and search neighborhood size are both set to $3 \times 3 \times 3$, and the regularization coefficients λ_1 and λ_2 are both set to 0.001.

3.1 Experiments on Simulated Dataset

In the simulation experiments, 10 subjects with manual hippocampal labels at year 1 are used as the baseline data (t=1), and then used to simulate atrophy on the hippocampi. Each subject is simulated with three longitudinal MR images, where images at year 2 (t=2) and year 3 (t=3) are generated by an atrophy simulation model [17] to ensure shrinking hippocampal volumes along the temporal dimension. Thus, ten sets of simulated longitudinal data with about 5% of annual hippocampal volume shrinking are obtained. The total simulated atrophy rate of hippocampus in three years is 9:15%.

A leave-one-out strategy on the total 10 simulated subjects is adopted to compare the segmentation performances of Non-local, SPBM, and our proposed method. Specifically, in each leave-one-out experiment, one subject is selected as 3D+t *subject* image, and the rest 9 subjects are selected as 3D *atlases* (27 *atlases* in total). For both Non-local and SPBM methods, three time-point images in the 3D+t subject image are segmented independently. The mean and standard deviation of Dice ratios are shown in Table 1. We can observe that our proposed method receives significant improvement over both Non-local and SPBM methods in terms of Dice ratio according to the paired *t*-test (p < 0.05). Thus, our proposed method achieves the best segmentation accuracy.

Figure 2(a) shows the curves of the longitudinal loss of overall hippocampus volume in a typical subject by Non-local, SPBM, and our method. Figure 2(b) shows the averaged loss of overall hippocampus volume estimated by these three methods. The final estimated mean and standard deviation of loss of overall hippocampus volume is $11:84\% \pm 1:63\%$ by Non-local, $12:31\% \pm 1:57\%$ by SPBM, and $10:87\% \pm 1:21\%$ by the proposed method. We can see that our proposed method is the closest to the ground truth, and also the most consistent in measuring longitudinal hippocampal volume changes, due to the use of temporal consistency constraint on sparse representation of multi-atlas based label fusion.

3.2 Experiments on Real Dataset

In the real experiments, we randomly select 12 subjects with each having 3 time points (baseline, 6 and 12 months) from ADNI dataset. The hippocampi of these MR images have been manually labeled, which are regarded as ground truth.

We also adopt a leave-one-out strategy on the total 12 real subjects for experiments of hippocampus segmentation. The mean and standard deviation of Dice ratios and also the average symmetric surface distance (ASSD) of the hippocampus segmentation results by Non-local, SPBM, and our proposed method are shown in Table 2. We can observe that our proposed method achieves significant improvement over Non-local and SPBM methods in terms of both Dice ratio and ASSD according to the pair *t*-test (p < 0.05). Figure 3 shows the surface distances of the left hippocampus for a typical subject, between manual segmentations and automatic segmentations by Non-local, SPBM, and our proposed 3D+t hippocampus segmentation method achieves the best segmentation performance.

4 Conclusion

In this paper, we proposed an integrated *temporal sparse representation* and *multi-atlas patch-based label fusion* method for longitudinal (3D+t) hippocampus segmentation in the longitudinal MR images. To make the registration at different time points consistent to the subsequent data analysis, we registered the 3D+t subject image and all atlases to the group-mean image of 3D+t subject image by using GLIRT. Moreover, to respect the smooth change of longitudinal structure (i.e., hippocampus), we added a temporal fused smoothness term to the objective function of sparse representation, for enforcing small difference between two successive sparse representation vectors from adjacent time points. Experimental results demonstrated the improved segmentation accuracy and longitudinal consistency by our proposed method, compared to both Non-local and SPBM methods.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China (No. 61503300) and China Postdoctoral Science Foundation (No. 2014M560801).

References

1. Bird CM, Burgess N. The hippocampus and memory: insights from spatial processing. Nat Rev Neurosci. 2008; 9(3):182–194. [PubMed: 18270514]

- Schuff N, et al. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain. 2009; 132(4):1067–1077. [PubMed: 19251758]
- Schröder J, Pantel J. Neuroimaging of hippocampal atrophy in early recognition of Alzheimer's disease - a critical appraisal after two decades of research. Psychiatry Res: Neuroimaging. 2016; 247:71–78. [PubMed: 26774855]
- 4. van der Lijn F, et al. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. NeuroImage. 2008; 43(4):708–720. [PubMed: 18761411]
- 5. Rousseau F, Habas PA, Studholme C. A supervised patch-based approach for human brain labeling. IEEE Trans Med Imaging. 2011; 30(10):1852–1862. [PubMed: 21606021]
- Zhang D, Guo Q, Wu G, Shen D. Sparse patch-based label fusion for multi-atlas segmentation. In: Yap PT, Liu T, Shen D, Westin CF, Shen L, editorsMBIA 2012 LNCS. Vol. 7509. Springer; Heidelberg: 2012. 94–102.
- 7. Zarpalas D, et al. Gradient-based reliability maps for ACM-based segmentation of hippocampus. IEEE Trans Biomed Eng. 2014; 61(4):1015–1026. [PubMed: 24658226]
- Song Y, Wu G, Sun Q, Bahrami K, Li C, Shen D. Progressive label fusion framework for multi-atlas segmentation by dictionary evolution. In: Navab N, Hornegger J, Wells WM, Frangi AF, editorsMICCAI 2015, Part III LNCS. Vol. 9351. Springer; Heidelberg: 2015. 190–197.
- 9. Wolz R, et al. Measurement of hippocampal atrophy using 4D graph-cut segmentation: application to ADNI. NeuroImage. 2010; 52(1):109–118. [PubMed: 20382238]
- Leung KK, et al. Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease. NeuroImage. 2010; 51(4):1345–1359. [PubMed: 20230901]
- Guo Y, Wu G, Yap PT, Jewells V, Lin W, Shen D. Segmentation of infant hippocampus using common feature representations learned for multimodal longitudinal data. In: Navab N, Hornegger J, Wells WM, Frangi AF, editorsMICCAI 2015, Part III LNCS. Vol. 9351. Springer; Heidelberg: 2015. 63–71.
- Chincarini A, et al. Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease. NeuroImage. 2016; 125:834–847. [PubMed: 26515904]
- 13. Wu G, Wang Q, Shen D. Registration of longitudinal brain image sequences with implicit template and spatial-temporal heuristics. NeuroImage. 2012; 59(1):404–421. [PubMed: 21820065]
- Jenkinson M, et al. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage. 2002; 17(2):825–841. [PubMed: 12377157]
- Vercauteren T, et al. Diffeomorphic demons: efficient non-parametric image registration. NeuroImage. 2009; 45(1):S61–S72. [PubMed: 19041946]
- Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci. 2009; 2(1):183–202.
- Karacali B, Davatzikos C. Simulation of tissue atrophy using a topology preserving transformation model. IEEE Trans Med Imaging. 2006; 25(5):649–652. [PubMed: 16689268]





Schematic diagram of the proposed 3D+t hippocampus segmentation method.



Fig. 2.

Demonstration of loss of overall hippocampus volume through 3 time points. (a) Loss of longitudinal hippocampus volume for a typical subject, and (b) average loss of longitudinal hippocampus volumes for all subjects.





Visualization of surface distances (in mm) for hippocampus segmentation results by three methods.

Table 1

Mean and standard deviation of Dice ratio in hippocampus segmentation on simulated atrophy data (Unit: %).

Method	Left hippo	Right hippo	Overall
Non-local	81.27 ± 1.98	80.08 ± 2.82	80.69 ± 2.12
SPBM	80.80 ± 1.71	79.55 ± 2.43	80.19 ± 1.76
Proposed	82.71 ± 1.51 *	81.86 ± 2.44 *	82.30 ± 1.76 *

* Indicates significant improvement over Non-local and SPBM methods (p < 0.05)

-

Table 2

Mean and standard deviation of Dice ratios and the average symmetric surface distance (ASSD) for automatic segmentations by Non-local, SPBM, and our proposed method.

	Method	Left hippo	Right hippo	Overall	
Dice ratio (%)	Non-local	83.53 ± 2.99	82.37 ± 4.02	82.93 ± 2.97	
	SPBM	82.32 ± 3.83	82.32 ± 2.64	82.31 ± 2.87	
	Proposed	84.85 ± 2.30 *	84.29 ± 2.34 *	84.55 ± 2.22 *	
ASSD (mm)	Non-local	0.497 ± 0.072	0.524 ± 0.098	0.512 ± 0.067	
	SPBM	0.520 ± 0.096	0.509 ± 0.057	0.515 ± 0.066	
	Proposed	$0.455 \pm 0.049^{\ast}$	$0.476 \pm 0.063^{\ast}$	0.466 ± 0.053 *	

* Indicates significant improvement over Non-local and SPBM methods (p < 0.05)