

Supervoxel-Based Hierarchical Markov Random Field Framework for Multi-Atlas Segmentation

Ning Yu¹, Hongzhi Wang², and Paul A. Yushkevich^{3*}

¹ Department of Computer Science, University of Virginia
85 Engineers Way, Charlottesville, VA 22904

ny4kt@virginia.edu

² IBM Almaden Research Center

650 Harry Road, San Jose, CA 95120

hongzhiw@us.ibm.com

³ Department of Radiology, University of Pennsylvania

3700 Hamilton Walk, Philadelphia, PA 19104

paully2@mail.med.upenn.edu

Abstract. Multi-atlas segmentation serves as an important technique for quantitative analysis of medical images. In many applications, top performing techniques rely on computationally expensive deformable registration to transfer labels from atlas images to the target image. We propose a more computationally efficient label transfer strategy that uses supervoxel matching regularized by Markov random field (MRF), followed by regional voxel-wise joint label fusion and a second MRF. We evaluate this hierarchical MRF framework for multi-label diencephalon segmentation from the MICCAI 2013 SATA Challenge. Our segmentation results are comparable to the top-tier one obtained by deformable registration, but with much lower computational complexity.

Keywords: Segmentation, Atlas, Supervoxel, MRF

1 Introduction

Accurate and efficient automated segmentation methods are highly sought in a range of biomedical imaging applications. The multi-atlas segmentation scheme, e.g. [10,18], has proven to be very accurate in a range of problems. Multi-atlas segmentation consists of two basic stages: *label transfer*, e.g., [8], and *label fusion*, e.g., [18].

Label transfer via affine registration to obtain global appearance matching is fast but has low accuracy in most applications. Label transfer via non-linear deformable registration, e.g., [3], leads to more accurate segmentation results by considering spatial variability of voxel-wise correspondence field, but at the cost of high computational burden. Recently, 3D supervoxels are introduced for cell segmentation by Lucchi et al. [12]. Supervoxel-based multi-atlas segmentation approaches [19,11] have eliminated the dependence of registration and have demonstrated comparative results in some segmentation applications. However, both kNN searching in [19] and the classification model in [11] ignore spatial relationships between supervoxels.

* This work was supported by the NIH grant R01 EB017255.

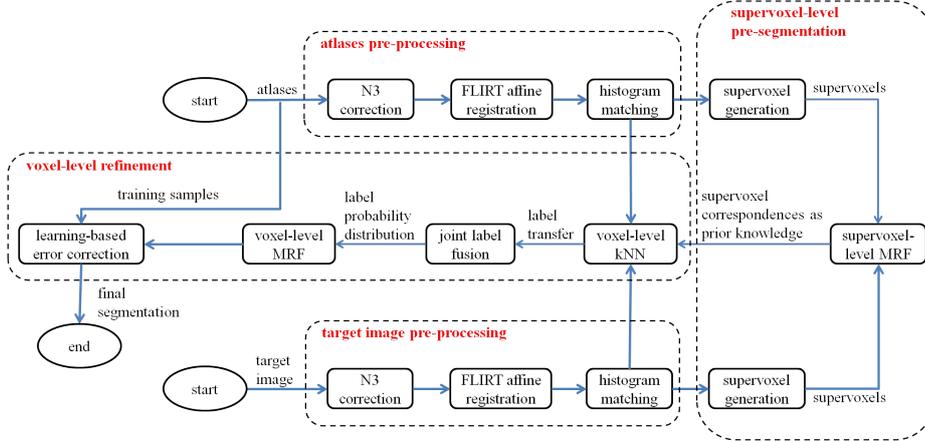


Fig. 1. Flowchart of the proposed framework. Three modules are indicated by the red titles.

In order to combine the individual strengths of 1) the efficiency of affine registration; 2) the accuracy of local supervoxel matching, we propose a supervoxel-based hierarchical Markov random field (MRF) framework for multi-atlas segmentation.

Our study is inspired by the work of [7], who reformulated the problem of dense deformable image registration as an MRF model. Some similar works include [9], where Heinrich et al. use the minimum-spanning-tree-based graphical model of overlapping layers of supervoxels to represent the image domain and then use belief propagation to solve the discrete optimization.

In the label fusion stage, patch-based methods [6,16] outperform global methods due to spatially variable weight assignments, which better compensate local registration errors. Especially, [16] leverages PatchMatch algorithm [4] to approximate and accelerate kNN searching, such that label fusion becomes close to real time.

However, these methods consider each atlas weight independently and may ignore their correlations in between. Additionally, performing label fusion at the supervoxel level may not adequately capture spatial variations for optimal label fusion [19]. As a result, we derive the hierarchical MRF framework, where joint label fusion [18] and learning-based error correction [17] are incorporated into the voxel-level refinement. Joint label fusion is used to minimize bias from correlated atlases, while error correction is used to learn and compensate the systematic segmentation error made by the wrapped framework. The flowchart in Fig.1 illustrates the framework.

2 Method

2.1 Pre-processing

All atlases and target images are pre-processed through the following pipeline: 1) inhomogeneity correction via N3 [14]; 2) affine registration to the ICBM152 template

We define the unary potential $D(\cdot, \cdot)$ as the Mahalanobis distance in feature space between target supervoxel and atlas supervoxel pointed by a given displacement vector: $D(l_P | f_P, A) = \text{dist}_{Mah}(f_P, f_{A(c_P + d^{l_P})})$, where the covariance matrix in the Mahalanobis distance metric is estimated by all the training features; c_P represents the center of target supervoxel P ; d^{l_P} represents the displacement vector indexed by the label l_P ; and $f_{A(c_P + d^{l_P})}$ represents the feature of the atlas supervoxel in A that covers the coordinate $c_P + d^{l_P}$.

We then define the pairwise potential $V(\cdot, \cdot)$ as the Manhattan distance in vector space, normalized by two factors, between the two displacement vectors indexed by the corresponding labels of two adjacent target supervoxels:

$$V(l_P, l_Q) = \frac{\tau_b(P, Q)}{\tau_d(P, Q)} \cdot \text{dist}_{Man}(d^{l_P}, d^{l_Q}), \quad (1)$$

where $\tau_b(P, Q) = |\{p \in P \mid \exists q \in Q, \text{ s.t. } q \text{ is adjacent to } p\}|$ is the boundary overlap factor of P and Q , i.e., the number of voxels at the boundaries of P and Q ; and $\tau_d(P, Q) = \|c_P - c_Q\|_2$ is the distance factor of P and Q , i.e., the Euclidean distance between the centers of P and Q . Fig.2 is a 2D diagram illustrating the formulation of the proposed supervoxel-based MRF model.

The proposed energy function is minimized by the Fast-PD algorithm used in [7]. After computing the correspondence field between supervoxels in the target image and in each of the atlases in the library, we assign to each target supervoxel the posterior probability of each anatomical label. This is done by simple majority voting.

2.3 Voxel-level refinement

Joint label fusion after kNN at voxel level We define the relevant supervoxels are those with the posterior probability of any anatomical label more than a specified threshold (e.g., 0.1 in our experiments). Then in the following steps, we do not need to consider those irrelevant supervoxels, which greatly narrows down the search domain. For each voxel p inside the relevant supervoxels, we search its k nearest neighbors as corresponding voxels from all the corresponding atlas supervoxel domains. k equals 20 in our experiments. The distance metric is defined as the Euclidean distance between the two normalized intensity vectors (elements with zero mean and one standard deviation) over the patches with size $5 \times 5 \times 5$ centered at the two voxels respectively. The labels of the corresponding atlas voxels are transferred and fused into a consensus label for the target voxel through the joint label fusion (JLF) technique [18].

MRF at voxel level JLF provides the probability vector v_p for each relevant target voxel p to belong to each label. Considering spatial smoothness, we apply the multi-labeling MRF model for the second time to satisfy voxel-wise labeling smoothness, where the label set \mathcal{L}' includes all relevant anatomical labels. Mathematically, voxels inside all the relevant target supervoxels construct a graph $\mathcal{G}'_T = \{\mathcal{N}'_T, \mathcal{E}'_T\}$, where the edges \mathcal{E}'_T are the four-connectivity lattice-like neighborhood system. Let \mathcal{L}'_T denote the set of all possible labels spanned for all nodes in \mathcal{N}'_T . The energy

Steps	Mean (Median) DSC	Mean time
N3 correction	-	4 min
FLIRT affine registration	-	7 min
Histogram matching	-	1 min
Supervoxel generation	-	4 min
Feature extraction	-	1 min
Supervoxel-level MRF	0.6336 (0.6462)	1 min
Voxel-level kNN + JLF + MRF	0.8399 (0.8462)	2 min
Learning-based ER	0.8548 (0.8625)	1 min
Final DSC and total time	0.8548 (0.8625)	21 min

Table 1. Means and medians of the average DSC by each step and means of the running time (per target image) for each step of the proposed framework over the training atlas library in the leave-one-out cross validation scheme.

function takes the form: $L^* = \arg \min_{L \in \mathcal{L}'_T} E(L|v, T, \mu) = \mu \sum_{p \in \mathcal{N}'_T} D'(l_p|v_p) + \sum_{\{p,q\} \in \mathcal{E}'_T} V'(l_p, l_q|T)$, where p and q are voxel indices and the coefficient μ controls the relative importance of the two potentials.

We define the unary potential negatively related to the JLF probability of a certain voxel to a certain label: $D'(l_p|v_p) = 1 - v_p(l_p)$.

We then define the pairwise potential negatively related to the absolute intensity difference of two adjacent voxels as long as they are assigned with different labels:

$$V'(l_p, l_q|T) = \begin{cases} 1 - \frac{|T_p - T_q|}{\max_{\{p',q'\} \in \mathcal{E}'_T} |T_{p'} - T_{q'}|}, & \text{if } l_p \neq l_q \\ 0, & \text{otherwise} \end{cases}. \quad (2)$$

MRF energy function at voxel level is minimized by the Fast-PD algorithm [7].

Learning-based error correction at voxel level As a final step, we correct systematic segmentation errors at voxel level by applying the learning-based error correction (ER) strategy [17] trained on the atlas library.

3 Experiments

We evaluate the proposed framework for multi-label diencephalon segmentation from MICCAI 2013 SATA Challenge [2], where there are 12 testing target images and 35 training atlases with 14 anatomical labels. The framework is implemented in C++ on a laptop with Intel i7 dual-core 2.40 GHz CPU and 8 G memory.

First, we test the proposed framework on the training atlas library based on the leave-one-out cross validation scheme. Accuracy is evaluated by the average Dice Similarity Coefficient (DSC) over all anatomical labels. Efficiency is evaluated by the running time per target image for each step. Results are summarized in Table 1. Here we set the optimal $\gamma = 3$ and $\mu = 2$ shown in the fine tuning curves (Fig.3), which are not particularly sensitive to the results.

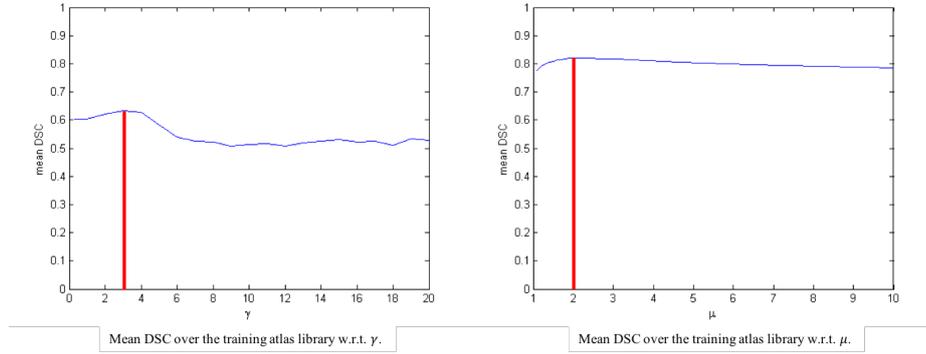


Fig. 3. Mean DSC over the training atlas library (in leave-one-out cross validation) with different settings of γ (Left) and μ (Right), respectively. Red lines indicate the optimal values which are also the settings in our experiments.

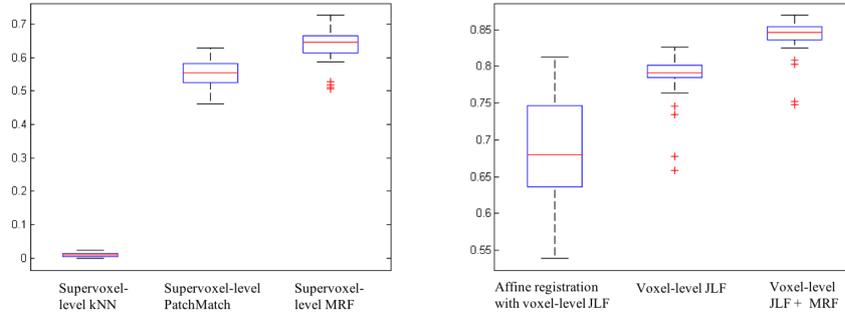


Fig. 4. Boxplots of DSC over the training atlas library. Left: By supervoxel-level kNN (left), supervoxel-level PatchMatch (middle), and supervoxel-level MRF (right), respectively. Right: By affine registration with voxel-level JLF (left), supervoxel-level pre-segmentation with voxel-level JLF (middle), and supervoxel-level pre-segmentation with voxel-level MRF (right), respectively.

Second, we evaluate the effectiveness of MRF at supervoxel level before voxel-level refinement. We compare the proposed supervoxel-level MRF with 1) the supervoxel-level kNN method [19], and 2) the PatchMatch scheme [4] at supervoxel level. Fig.4 Left shows the boxplots of DSC over the training atlas library by the three methods. Supervoxel-level kNN completely fails on our dataset (mean DSC of 0.0093). The proposed supervoxel-level MRF obtains pre-segmentation results (mean DSC of 0.6336) that largely exceed those of supervoxel-level PatchMatch (mean DSC of 0.5516).

Third, we evaluate the effectiveness of MRF at voxel level. We compare the voxel-level JLF segmentation with and without MRF. We also compare with the baseline segmentation provided by affine registration with voxel-level JLF. Fig.4 Right shows the boxplots of DSC over the training atlas library by the three methods. The proposed voxel-level MRF generates better segmentation results (mean DSC of 0.8399)

Methods	Mean (Median) DSC	Mean time
ANTs SyN [3] + JLF [18] + ER [17]	0.8663 (0.8786)	> 1 h
Proposed framework	0.8343 (0.8421)	21 min
Atlas Forest [20]	0.8282 (0.8484)	20-25 min

Table 2. Means and medians of the average DSC and means of running time (per target image) by each method over the testing images. The results are reported in [2].

than those of voxel-level JLF without MRF (mean DSC of 0.7844) and those of affine registration with JLF (mean DSC of 0.6828).

Additionally, over the testing images, we compare the proposed framework with one top-tier multi-atlas segmentation method [3] which includes the pairwise deformable registration. We also involve the random-forests-based atlas coding (Atlas Forest) [20], as another efficient method with single deformable registration (instead of several atlas-target pairwise registrations), for comparison. The implementation details and results of the compared methods are reported in [2]. Table 2 summarizes the overall performance on the same dataset. In general, our framework reaches better trade-off between accuracy and time complexity. Note that: 1) for practice consideration we use a global template for affine registration rather than perform atlas-target pairwise registration, which reduces the complexity by a factor of the atlas library size (35 in our experiment), and 2) the theoretical efficiency gain of the proposed framework is much higher than the "wall" time gain because the complexity of the registration problem over dense deformable registration is reduced by a factor of the supervoxel size (~ 125 in our experiment). Most of our time is spent in pre-processing and graph construction stages. **The actual time spent in FastPD is less than 5 seconds per target-atlas pair.**

4 Conclusion

In this paper, we proposed a supervoxel-based hierarchical MRF framework for efficient multi-atlas segmentation. In experimental validation, we showed encouraging performance: at the same rate of segmentation accuracy, our framework is much more efficient than one of top-tier state-of-the-art methods, enabling it more practical on large datasets or high-dimension images for clinical analysis.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34(11), 2274–2282 (2012)
2. Asman, A., Akhondi-Asl, A., Wang, H., Tustison, N., Avants, B., Warfield, S.K., Landman, B.: Miccai 2013 segmentation algorithms, theory and applications (sata) challenge results summary. In: *MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications (SATA)* (2013)
3. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis* 12(1), 26–41 (2008)

4. Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.B.: PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 28(3) (2009)
5. Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C.: Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. *Journal of computer assisted tomography* 18(2), 192–205 (1994)
6. 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(2), 940–954 (2011)
7. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through mrfs and efficient linear programming. *Medical image analysis* 12(6), 731–741 (2008)
8. Glocker, B., Sotiras, A., Komodakis, N., Paragios, N.: Deformable medical image registration: Setting the state of the art with discrete methods*. *Annual review of biomedical engineering* 13, 219–244 (2011)
9. Heinrich, M.P., Simpson, I.J., Papież, B.W., Brady, M., Schnabel, J.A.: Deformable image registration by combining uncertainty estimates from supervoxel belief propagation. *Medical image analysis* 27, 57–71 (2016)
10. Iglesias, J.E., Sabuncu, M.R.: Multi-atlas segmentation of biomedical images: a survey. *Medical image analysis* 24(1), 205–219 (2015)
11. Kanavati, F., Tong, T., Misawa, K., Fujiwara, M., Mori, K., Rueckert, D., Glocker, B.: Supervoxel classification forests for estimating pairwise image correspondences. In: *Machine Learning in Medical Imaging*, pp. 94–101. Springer (2015)
12. Lucchi, A., Smith, K., Achanta, R., Knott, G., Fua, P.: Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features. *IEEE transactions on medical imaging* 31(2), 474–486 (2012)
13. Rolland, J.P., Vo, V., Abbey, C., Bloss, B.: Fast algorithms for histogram matching: Application to texture synthesis. *Journal of Electronic Imaging* 9(1), 39–45 (2000)
14. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in mri data. *Medical Imaging, IEEE Transactions on* 17(1), 87–97 (1998)
15. Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al.: Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage* 23, S208–S219 (2004)
16. Ta, V.T., Giraud, R., Collins, D.L., Coupé, P.: Optimized patchmatch for near real time and accurate label fusion. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 105–112. Springer (2014)
17. Wang, H., Das, S.R., Suh, J.W., Altinay, M., Pluta, J., Craige, C., Avants, B., Yushkevich, P.A., Initiative, A.D.N., et al.: A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. *NeuroImage* 55(3), 968–985 (2011)
18. Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A.: Multi-atlas segmentation with joint label fusion. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35(3), 611–623 (2013)
19. Wang, H., Yushkevich, P.A.: Multi-atlas segmentation without registration: A supervoxel-based approach. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pp. 535–542. Springer (2013)
20. Zikic, D., Glocker, B., Criminisi, A.: Atlas encoding by randomized forests for efficient label propagation. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pp. 66–73. Springer (2013)