Globally Optimal Label Fusion
with Shape Priors

Ipek Oguz[1,2(✉)], Satyananda Kashyap[2], Hongzhi Wang[3], Paul Yushkevich[1],
and Milan Sonka[2]

[1] Department of Radiology, University of Pennsylvania, Philadelphia, USA
`ipekoguz@mail.med.upenn.edu`
[2] Iowa Institute for Biomedical Imaging, University of Iowa, Iowa City, USA
[3] IBM Research, Almaden, USA

Abstract. Multi-atlas label fusion methods have gained popularity in a variety of segmentation tasks given their attractive performance. Graph-based segmentation methods are widely used given their global optimality guarantee. We propose a novel approach, GOLF, that combines the strengths of these two approaches. GOLF incorporates shape priors to the label-fusion problem and provides a globally optimal solution even for the multi-label scenario, while also leveraging the highly accurate posterior maps from a multi-atlas label fusion approach. We demonstrate GOLF for the joint segmentation of the left and right pairs of caudate, putamen, globus pallidus and nucleus accumbens. Compared to the FreeSurfer and FIRST approaches, GOLF is significantly more accurate on all reported indices for all 8 structures. We also present comparisons to a multi-atlas approach, which reveals further insights on the contributions of the different components of the proposed framework.

1 Introduction

Automated quantification of basal ganglia from MRI scans is of great importance for many neuroimaging studies. For example, striatal atrophy, assessed via volume measurements of the caudate and the putamen, is the single most prominent imaging-based measure of Huntington's Disease [8]. Despite their importance, some subcortical structures, such as the nucleus accumbens (NA) and the globus pallidus (GP), are notoriously difficult to segment robustly because the MRI signal often lacks enough contrast to distinguish them from neighboring structures. This leads to substantial measurement noise and has been a barrier to the robust discovery of potential abnormalities in these structures.

Existing subcortical segmentation methods include probabilistic approaches [4] and deformable shape models [7]. These methods typically do not consider interactions between these neighboring structures. Multi-atlas label fusion (MALF) approaches [5] have recently gained popularity as they have been consistently the best performers for many challenging segmentation tasks, including subcortical segmentation. The general MALF idea is to register multiple annotated atlases to the subject image, and to combine these alternative solutions into a coherent segmentation. However, these methods typically do not provide

a way to incorporate prior knowledge of shape, with the notable exception of [9] where a contour-driven prior distribution is used along with image features. Another challenge is the difficult-to-register features present in the subcortical structures, such as the thin tail of the caudate.

Graph-cuts have been recently proposed [6] as a novel way for optimization in the MALF framework, but the proposed method again did not allow shape priors. Furthermore, obtaining the globally optimal solution was only guaranteed for binary labeling, and approximate solutions are needed for multi-label segmentation. This is in general true of image-grid based graph cuts [2], where approximations such as α-expansions and α-β swaps [3] were proposed for the multi-label case. In contrast, the surface-based LOGISMOS [11] framework sustains the optimality guarantee even in the multi-label scenario, but has limitations on the types of inter-object relations that can be represented (Sect. 2.4). LOGISMOS object-to-object mapping is also problematic for many applications including high-curvature objects and overlapping initial segmentations (Sect. 2.3).

We propose a new graph-based MALF approach that overcomes the challenges of the LOGISMOS framework, incorporates shape priors to the label-fusion problem and provides a solution that is globally optimal with respect to the provided cost functions and constraints, even for the multi-label scenario. We apply this new approach named GOLF (Globally Optimal Label Fusion) to the joint segmentation of the caudate nucleus, putamen, GP and NA.

Our main contributions are: (1) Discrete optimization framework for MALF segmentation. (2) Introducing shape priors to the MALF problem, which enforces more appropriate regularization than currently possible. (3) Globally optimal solution guarantee even for multi-label problems. (4) Generalization of LOGISMOS to arbitrary inter-object relationships. (5) Novel object-to-object mapping method for graph construction. (6) Application to subcortical segmentation.

2   Methods

The GOLF approach transforms the optimal image segmentation task to a max-flow problem on a geometric graph. The node costs are given by the posterior maps from a multi-atlas labeling approach (Sect. 2.1). Each object (label) is represented by a geometric sub-graph that incorporates a shape prior and smoothness constraints (Sect. 2.2). Multi-object segmentation is achieved by combining these sub-graphs into a larger graph, which requires object-to-object mapping (Sect. 2.3). The inter-object arcs allow the representation of constraints between neighboring objects and can represent arbitrary relationships, unlike the original LOGISMOS (Sect. 2.4). After graph construction, an s-t cut algorithm is used to jointly optimize the segmentation of all the objects in the graph [11].

2.1   Joint Label Fusion

We begin by performing multi-atlas segmentation using the joint label fusion (JLF) approach described in [10]. Briefly, each atlas is first deformably registered to the testing image. For each voxel, the warped labels from each atlas

constitutes a "vote". A popular strategy is to weight these votes based on the appearance similarity of a small neighborhood, which provides an estimate of the local segmentation quality for a particular atlas. However, this strategy estimates voting weights for each atlas independently, which fails to take into account bias from similar atlases. Instead, the JLF approach estimates a matrix of pairwise atlas dependencies, which alleviates this bias. Given a testing image, JLF produces a labeling of the image and posterior maps for each label.

2.2   Graph Construction for Binary Segmentation

A mean shape model is constructed from the training set for each object (i.e., for the left/right caudate, putamen, GP, NA) following rigid registration. The model is obtained by averaging the registered binary models and thresholding voxels with 25 % or more agreement. A mesh model is obtained using marching cubes and smoothing. Each testing image was initially segmented with the JLF method (Sect. 2.1). The voxels segmented as one of the objects of interest were extracted, and the mean shape model was fitted to these initial segmentations using an iterative closest points approach. While the JLF approach provides excellent localization of each structure, the segmentation accuracy may be inadequate, especially in difficult-to-register areas such as the thin tail of the caudate, and may contain holes or handles. Fitting the shape model provides a more appropriate representation of the entire structure including its fine-level features, and ensures accurate topology.

After model fitting, at each vertex of this initial surface, a "column" consisting of a set of graph nodes placed at increasing distances from the surface is built. The column represents the set of candidate locations for the final surface, i.e., the local search space, as in the original LOGISMOS approach [11]. Each node has an associated cost given by the JLF posterior at that location. The LOGISMOS approach transforms the surface segmentation problem into the minimum closed set problem. A closed set C in a graph X is a subgraph such that all successors in X of nodes in C are also in C. The cost of a closed set is the total cost of the nodes in the set. The minimum closed set of a graph X can be identified in polynomial time by computing a minimum s-t cut in a graph derived by introducing intra-column arcs between successive nodes in each column. Inter-column arcs enforce smoothness constraints. The equivalence to the minimum closed set is accomplished by transforming the node weights (Fig. 1). The shape prior is encoded by the initial shape, which limits the search space.

An important concern is that the paths of the columns should not cross each other, since crossing columns can lead to topological defects such as self-intersections in the final surfaces. We use the non-intersecting electric lines of force (ELF) [11] concept for this purpose. We simulate placing a positive charge at each vertex of the mesh surface and compute the electric field these particles form, using Coulomb's law; the graph columns follow the electric lines of force.
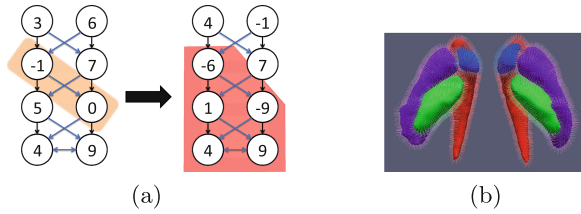
Fig. 1. (a) LOGISMOS graph structure. Nodes are organized into columns. The graph is transformed by subtracting each node's cost from the node above it. Intra-column arcs (black) ensure the cost of the optimal surface (orange) is equal to the cost of the min closed set (red). Inter-column arcs (blue) add smoothness constraints. **(b)** The entire graph for a representative subject.

2.3   Multi-Label Segmentation: Object-to-object Mapping

The multi-object segmentation is accomplished in the LOGISMOS framework by linking the graphs representing the interacting objects via inter-object arcs between columns of the two objects. These infinite-weight arcs enforce separation constraints between the objects. However, finding a suitable one-to-one mapping between the two objects is not a trivial task. Intersecting paths in this mapping can result in topological defects on the final segmentation.

   The original LOGISMOS framework attempts to solve this problem by detecting the medial surface between the two objects in a region of interaction (ROI). Starting from the first object, the graph column is "pushed forward" along the first object's ELF field until the medial surface, then is "pulled backward" along the second object's ELF field (Fig. 2-a, b). While this approach may be adequate for low-curvature surfaces, it is inappropriate for higher-curvature objects since the two ELF fields may not be properly aligned, which results in sharp artificial "corners" in the graph columns. This approach is also prone to discretization errors and has an additional computational cost for the medial surface detection; furthermore, the need for a pre-determined ROI introduces additional singularities in the graph structure along the border of the ROI.

   GOLF adopts a novel object-to-object mapping strategy to resolve these problems. Instead of computing disjoint ELF fields and transitioning from ELF to ELF, we build a single, coherent ELF field. This is accomplished by putting a positive charge on the vertices of the current object and negative charges on the vertices of all other objects (Fig. 2-c). Simple physics rules ensure that the lines of force (and therefore the graph columns) "flow" from the current object to neighboring objects in an orderly fashion, without intersections. This removes the need for the computation of either a medial surface or an arbitrary ROI, since the effect of the charged particles naturally decline with distance. This results in a single, seamless ELF field that allows for a better graph construction. The joint ELF computation also allows for overlapping initial segmentations (Fig. 2-d), which can often happen during the shape model fitting stage. While it would be possible to mask the initial segmentations to avoid overlaps, such an approach
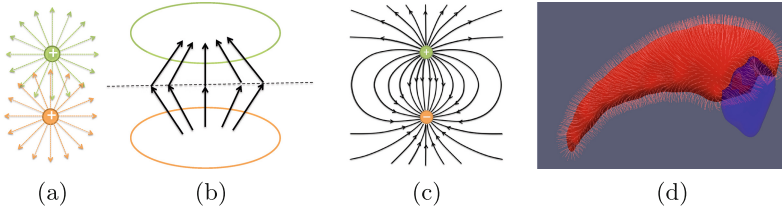
Fig. 2. (a) LOGISMOS object-to-object mapping. Every object has its own, disjoint ELF field. **(b)** LOGISMOS column construction. The medial surface (dashed) is computed; the column is pushed forward along the ELF of the first object until the medial surface, and then pulled back along the ELF of the second object. The transition creates corners and singularities. **(c)** GOLF's joint ELF field provides smooth search paths. **(d)** GOLF graph construction seamlessly handles overlapping initial segmentations.
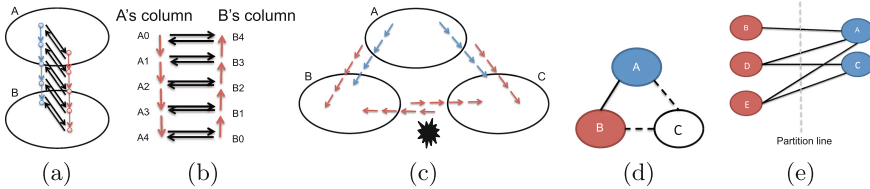

Fig. 3. (a) Columns must be reversed between interacting objects. Blue: A's column; red: B's column; black: inter-object arcs. **(b)** Linking columns with same orientation would result in trivial cuts. **(c)** This requirement is problematic for 3 mutually interacting objects. **(d)** Column ordering as a 2-coloring problem. Object C can not be colored. **(e)** LOGISMOS is limited to inter-object relationships represented by bipartite graphs.

is undesirable as it would have distorted the imposed shape prior. In contrast, the proposed approach seamlessly handles overlapping initial segmentations.

2.4 Multi-Label Segmentation: Arbitrary Inter-object Relationships

To combine the subgraphs for each object into a single graph for joint optimization, the LOGISMOS framework requires that the column direction between two interacting objects be of opposite polarity (inside-out vs. outside-in, Fig. 3-a). Without this requirement, every feasible cut through the graph is an empty cut (Fig. 3-b). While this is a harmless requirement in the case of 2 objects, it becomes problematic for larger sets of mutually interacting objects (Fig. 3-c).

Let us consider a graphic representation of the inter-object relationships such that each object is represented by a node. The two possible graph column orientations can be identified as two possible colorings (red and blue) of the corresponding nodes (Fig. 3-d). By the requirement above, inter-object arcs can only be inserted between objects that have different colors, i.e., opposite column directions. This problem is equivalent to the classical graph coloring problem with $n = 2$ allowed colors. It is well known that the only graphs that can be 2-colored are bipartite graphs (Fig. 3-e).
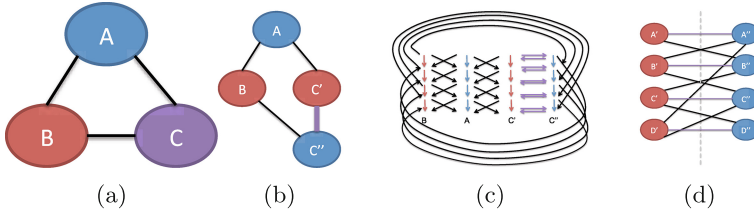
Fig. 4. (a, b) In GOLF, objects that have conflicting coloring requirements are replaced with two new identical subgraphs. **(c)** Arcs in the new GOLF hierarchy. Red, blue: representative columns from red and blue objects, respectively. Black: inter-object arcs. Purple: the "purple" arcs that guarantee coherent segmentation of the pseudo-objects C' and C". **(d)** GOLF can represent arbitrary inter-object relationships.

GOLF provides a novel graph construction method that allows non-bipartite inter-object relationships. Let's revisit the example in Fig. 3-d, where the object C has conflicting coloring requirements. We remove the object C and replace it with two identical objects, C' and C" (Fig. 4-a, b). The subgraphs representing C' and C" are both identical to the original subgraph C in terms of nodes, arcs and columns. Between each node in C' and each corresponding node in C", we insert a pair of special "purple" arcs with infinite weight, which ensures the cut through C' and C" is coherent (Fig. 4-c). The node cost for each node in C' and C" are half of the cost of the corresponding node in the original subgraph C, such that the total cost of a cut through C' and C" is equivalent to the cost of a cut through the original C. C' is colored red and thus can interact with A; C" is blue and can interact with B. This construction allows the representation of the mutual relationships between A, B, and C, which was previously impossible. To generalize, each object can be colored "purple" in this manner, such that any set of inter-object relationships can be represented (Fig. 4-d).

3   Experimental Methods

We used the public MICCAI Multi-Atlas Labeling Challenge dataset[1], which contains T1w brain images and expert manual segmentations for 35 subjects (ages 18–90). As in the Challenge, we used 15 images as atlases and 20 images for testing. The following automated segmentation methods were compared for the left/right pairs of caudate, putamen, GP and NA: (1) The proposed GOLF approach. (2) Majority voting (MV), as implemented in the ITK library, to assess the improvement of the proposed method over a simplistic multi-atlas labeling approach. 3) FreeSurfer [4], version 5.1. 4) FIRST, which is part of FSL v6.0 [7]. For GOLF, each column consisted of 30 nodes at 0.15 mm intervals. Inter-column smoothness constraints were set to 1 node interval. Object-specific settings are listed in Fig. 5-d. For each method, we report Dice and Jaccard overlap and

---

[1] https://masi.vuse.vanderbilt.edu/workshop2012/index.php.

surface-to-surface distances between the automated results and the expert tracings. For all volumetric segmentations, marching cubes and smoothing were used to generate surface representations for surface-based error computation. Paired t-tests are used for statistical comparison with a significance threshold of 0.05.

4   Results

Figure 5 presents the quantitative comparison of the results. Compared to the FS and FIRST methods, the proposed GOLF approach performed significantly better ($p \ll 0.001$) on all reported measures. The FIRST software only produced results on 11 out of the 20 test images. Compared to MV, GOLF performed significantly better ($p < 0.01$) on all reported measures for the caudate. For the NA, the improvements were significant only for the surface-based measurements ($p < 0.01$) but not for Dice and Jaccard coefficients. The differences between MV and GOLF did not reach significance for putamen and GP, with the exception of the surface error for the left putamen and left GP.

These findings suggest that the multi-atlas approach for cost estimation, the shape prior, and the graph-based optimization all contribute to the improvements in accuracy obtained with the GOLF approach. In particular, we note that GOLF is significantly more accurate for the caudate compared to the MV approach. The improvement in the surface-based measurements for the caudate are more dramatic than the volumetric measures, suggesting the thin tail of the caudate (which doesn't heavily contribute to the overall volume) is better captured by leveraging the shape prior in the GOLF framework. This is also supported by the surface errors obtained by the JLF approach alone (without the graph), 0.55 and 0.61 mm for the left and right caudate respectively, indicating that both the multi-atlas and the graph components are substantially contributing to the observed improvements in GOLF. In contrast, for putamen and GP, which are more blob-like in geometry, there is little to no improvement compared to the MV approach, suggesting the multi-atlas approach is the driving factor. Further analysis of the individual contributions of these components remains as future work.

Dice	FS	FIRST	MV	GOLF
N	20	11	20	20
L-accumbens	0.615	0.690	0.777	0.776
L-caudate	0.807	0.791	0.805	0.872
L-pallidus	0.767	0.793	0.867	0.869
L-putamen	0.777	0.844	0.911	0.910
R-accumbens	0.452	0.658	0.759	0.773
R-caudate	0.788	0.839	0.792	0.872
R-pallidus	0.802	0.779	0.871	0.869
R-putamen	0.792	0.862	0.909	0.906

(a)

Jaccard	FS	FIRST	MV	GOLF
N	20	11	20	20
L-accumbens	0.447	0.533	0.641	0.637
L-caudate	0.681	0.678	0.688	0.778
L-pallidus	0.630	0.663	0.765	0.769
L-putamen	0.636	0.740	0.838	0.836
R-accumbens	0.294	0.497	0.618	0.633
R-caudate	0.652	0.726	0.669	0.778
R-pallidus	0.675	0.647	0.774	0.772
R-putamen	0.657	0.764	0.834	0.831

(b)

Surface-to-surface error	FS	FIRST	MV	GOLF
N	20	11	20	20
L-accumbens	1.16	0.74	0.57	0.47
L-caudate	1.19	0.76	0.86	0.49
L-pallidus	0.82	0.77	0.42	0.52
L-putamen	1.29	0.60	0.33	0.43
R-accumbens	1.69	0.71	0.59	0.51
R-caudate	1.39	0.69	0.95	0.49
R-pallidus	0.71	0.87	0.40	0.42
R-putamen	1.17	0.54	0.34	0.38

(c)

Number of vertices	Left	Right
Caudate	2998	3058
Putamen	3432	3312
Pallidus	1414	1422
Accumbens	830	756

(d)

Fig. 5. Segmentation accuracy. **(a)** Dice coefficient. **(b)** Jaccard coefficient. **(c)** Surface-to-surface error. **(d)** Object-specific parameter settings used for GOLF.

5   Discussion

The proposed GOLF approach combines the strengths of the graph-based optimization techniques and multi-atlas label fusion techniques. While the graph-based optimization scheme guarantees obtaining the globally optimal solution given the objective cost function and set of constraints, the choice of the objective function itself is equally crucial. Here, we leverage the JLF approach to obtain a suitable cost function. The presented results illustrate the superior performance of our approach compared to both other multi-atlas methods as well as currently popular software suites for subcortical segmentation. Validation on pathological datasets remain as future work. While the experimental design has not included other popular label fusion methods such as STAPLE, previous studies have shown that JLF out-performs these methods in various datasets [10], which is why it was chosen for our graph formulation. We note that for some applications, fuzzy segmentation approaches can be attractive to potentially better capture the inherent uncertainty in medical images, while others require discrete segmentations for subsequent analysis. Finally, another related approach is the formulation of multi-atlas segmentation as a nonparametric regression problem to estimate the expected error as a function of the number of atlases [1].

The technical contributions of the paper include addressing several of the shortcomings of the original LOGISMOS framework for multi-object segmentation, including object-to-object mapping and generalization to arbitrary sets of inter-object relationships. While the present paper focuses on the subcortical segmentation, the framework is directly applicable to other segmentation tasks.

References

1. Awate, S.P., Whitaker, R.T.: Multiatlas segmentation as nonparametric regression. IEEE TMI **33**(9), 1803–1817 (2014)
2. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient N-D image segmentation. Int. J. Comput. Vis. **70**(2), 109–131 (2006)
3. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. In: ICCV, pp. 377–384 (1999)
4. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron **33**(3), 341–355 (2002)
5. Iglesias, J.E., Sabuncu, M.R.: Multi-atlas segmentation of biomedical images: A survey. Med. Image Anal. **24**(1), 205–219 (2015)
6. Koch, L.M., Rajchl, M., Tong, T., Passerat-Palmbach, J., Aljabar, P., Rueckert, D.: Multi-atlas segmentation as a graph labelling problem: application to partially annotated atlas data. In: Ourselin, S., Alexander, D.C., Westin, C.-F., Cardoso, M.J. (eds.) IPMI 2015. LNCS, vol. 9123, pp. 221–232. Springer, Heidelberg (2015). doi:10.1007/978-3-319-19992-4_17

7. Patenaude, B., Smith, S., Kennedy, D., Jenkinson, M.: A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage **56**(3), 907–922 (2011)
8. Paulsen, J.S., Long, J.D., Johnson, H.J.: Clinical and biomarker changes in premanifest HD show trial feasibility: a decade of the PREDICT-HD study. Front. Aging **6**, 78 (2014)
9. Wachinger, C., Sharp, G.C., Golland, P.: Contour-driven regression for label inference in atlas-based segmentation. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8151, pp. 211–218. Springer, Heidelberg (2013). doi:10.1007/978-3-642-40760-4_27
10. Wang, H., Suh, J.W., Das, S.R., Pluta, J., Craige, C., Yushkevich, P.A.: Multi-Atlas segmentation with joint label fusion. IEEE PAMI **35**(3), 611–623 (2012)
11. Yin, Y., Zhang, X., Williams, R., Wu, X., Anderson, D.D., Sonka, M.: LOGISMOS. IEEE Trans. Med. Imaging **29**(12), 2023–2037 (2010)