

# Topology-Preserving STAPLE

John A. Bogovic, Pierre-Louis Bazin, Jerry L. Prince  
Johns Hopkins University, Baltimore, MD 21218

## Abstract

*Methodology for fusing multiple segmentations to produce an improved result has been useful in computational anatomical studies. Although obtaining segmentations of anatomy having a particular topology are essential to studies using diffeomorphic deformation based analyses, no methods of label fusion presented to date have incorporated information regarding the topology of the anatomy. In this paper, we introduce “Topology STAPLE”, a novel method that statistically fuses multiple rater segmentations into a topologically correct segmentation. We evaluate the method on both simulated data and real delineations of the cerebellum produced by human raters.*

## 1. Introduction

The computational study of anatomy [8, 17] has flourished since the advent of high-resolution magnetic resonance imaging (MRI). The variability in the size and shape of different anatomical structures has been studied in both normal and diseased populations [15, 13]. These studies rely on the accurate and precise delineations of the structures of interest. Delineations are often done by human raters, but have also been produced by automatic and semi-automatic segmentation algorithms.

The delineation of anatomy is vital to methods involving deformation-based morphometry (DBM) [1], statistical shape analysis [6], and atlas construction using diffeomorphic mappings [10]. DBM type analyses can localize differences in the size and shape within the structure of interest. Specifically, DBM methods reveal these local differences by studying the deformations that map each instance of that anatomy to a common, stereotactic space. Statistical shape analyses also study the differences in shape, but do so by studying the differences in shapes that result after a coarse registration. Diffeomorphic mappings have been widely sought after in the field of registration and atlas construction because they have the property that a one-to-one correspondence exists between points in the original and deformed space. This is desirable because it gives a one-to-one mapping from one subject’s anatomy to any other sub-

ject’s anatomy. One can equivalently say that the anatomy of every person is topologically equivalent. Naturally, the stated assumption may be invalid for structures that are complex and highly-varying between individuals (e.g., vascular structures), in which case these methods should be avoided. Nevertheless, when applicable, DBM and statistical shape analysis also operate under this supposition.

The pervasiveness of this assumption in many analysis methods underscores the importance of ensuring that the delineations of anatomy produced by human raters or automated methods have a particular topology. In fact, the one-to-one mapping that is assumed *does not exist* when the objects under consideration differ topologically. Automated segmentation methods that enforce topology have been proposed [12, 3], but human raters can only be encouraged to produce smooth results as part of a manual protocol. To our knowledge, no manual protocols for the delineation of structures mention object topology. In fact, topological defects are likely to be observed when analyzing human drawn segmentations due to the challenges of delineating highly complex and convoluted structures in three dimensions, especially given the requirement of working with one two-dimensional slice at a time.

A number of methods have been introduced to estimate both rater performance and the true underlying parcellation [18, 14]. It has been shown that the delineation produced by combining the results of raters is more reliable than a single rater’s result [14]. Recently, multiple atlas registration methods have proven to be an efficient tool for segmentation, when many atlases are combined [11]. While these methods usually include a mechanism for encouraging smoothness in the resulting segmentation, such as a Markov random field, none explicitly enforce a particular topology. In fact, enforcing a particular topology yields smooth labels inside of objects, without oversmoothing object boundaries and is especially important when dealing with objects with boundaries of high curvature.

In this paper, we describe “Topology STAPLE”, a method for estimating rater performances and the true, topologically correct, underlying segmentation. The details of our method are described in section 2, and include a brief summary of the STAPLE algorithm [18] on which

it is based. Our method is then tested on simple, simulated data to demonstrate the effectiveness of the algorithm, the results of which are described in section 3. We then applied our method to a set of 22 real, hand-drawn delineations of the human cerebellum to illustrate its relevance and applicability in the context of a research or clinical study.

## 2. Methods

In this section we describe our methodology, which follows the framework established in the STAPLE algorithm to produce a maximum-likelihood estimate of the rater performance parameters. It is a special case of the expectation-maximization (EM) algorithm, and estimates the fixed but unknown true segmentation (the “hidden” data) as well. Our methodology is nearly identical in formulation and similar in execution, adding an important topological constraint on the unknown true segmentation. We first introduce our notation.

We denote the rater data as  $\mathbf{D}_i = D_{ir}$ , where  $i \in \{1, 2, \dots, N\}$  indexes the voxels and  $r \in \{1, 2, \dots, R\}$  indexes the raters, where  $R$  is the number of raters. In this paper, theory is described only for binary labelings, implying that  $D_{ir} \in \{0, 1\}$ , although the method can be generalized as demonstrated in section 3.3. The performance of the  $r^{\text{th}}$  rater is parameterized by the true positive and true negative rates, denoted  $\theta_{r11}$  and  $\theta_{r00}$ , respectively. The perfect rater will therefore be parameterized by  $\theta_{r00} = \theta_{r11} = 1$ , (i.e., the rater’s true positive and true negative rates equal one). This subscript notation was chosen to emphasize that these parameters may be interpreted as the conditional probabilities:  $\theta_{r11} = p(D_{ir} = 1 | T_i = 1)$  and  $\theta_{r00} = p(D_{ir} = 0 | T_i = 0)$  where  $T_i$  denotes the true label at voxel  $i$ . The probabilities of error are therefore  $\theta_{r10} = 1 - \theta_{r00} = p(D_{ir} = 1 | T_i = 0)$  and  $\theta_{r01} = 1 - \theta_{r11} = p(D_{ir} = 0 | T_i = 1)$ . The performance of all raters are summarized by the  $2R$  vector  $\Theta = \theta_{rss}$ , where  $s \in \{0, 1\}$ .

Maximum-likelihood estimation of the rater parameters is performed using the expectation-maximization (EM) algorithm. The E-Step is performed as in STAPLE and estimates the underlying, true segmentation. Next, this segmentation is projected onto the topologically correct space using a topology correction algorithm [2]. Finally, the ML estimates of the rater performance are computed in the M-Step. This procedure is described in detail below.

### 2.1. E-Step

The E-Step of the EM iterations estimates the probability that the object is present at each voxel, and is computed by taking the conditional expectation of the complete data log likelihood,  $p(\Theta, \mathbf{T})$ , where  $\mathbf{T} = T_i$  gives the true label at each voxel  $i$ . After some manipulation, weight variables for

the iteration  $k$ ,  $W_{si}^{(k)}$  can be computed using:

$$W_{si}^{(k)} = p(T_i = s | \mathbf{D}_i, \Theta^{(k)}) \quad (1)$$

$$= \frac{p(T_i = s) \prod_{r: D_{ir}=s'} \theta_{rss'}^{(k)}}{\sum_s p(T_i = s) \prod_{r: D_{ir}=s'} \theta_{rss'}^{(k)}}. \quad (2)$$

the details of which can be found in [18] or [14] and where  $s \in \{0, 1\}$  as above. The result is a measure of the conditional probability that the object is present at each voxel  $i$  given the rater data  $\mathbf{D}_i$  and the estimated rater performances  $\Theta$ . These probability maps are likely to be topologically defective, as the conditional probabilities are measured locally. A Markov random field (MRF) can be applied to encourage homogeneity, but does not guarantee that a particular topology results. To remedy this, we apply a topology correction algorithm on the probability maps.

### 2.2. Topology Correction

We posit that the topology of the true, underlying object is known a priori and enforce that topology on our estimate of the segmentation. A topology correction algorithm using a fast-marching method [2] is applied directly to the object probability map computed as part of the E-Step. The result of this correction is a new probability map that approximates the original, but for which all isocontours or isosurfaces have the desired topology. Forming a hard segmentation of an object by thresholding this corrected probability map at any isovalue will result in an object with the correct topology. Before delving into the details of the methodology, we review some basics of digital topology.

Elementary topological changes to discrete objects involve the addition or removal of a point from the object. Points that do not affect the object’s topology when added or removed are called “simple points.” In fact, whether a point is simple or not may be determined from the point’s local neighborhood [4]. In this application, we seek to correct the topology of probability maps, and so make use of generalized simple points for scalar fields [5]. For a scalar field, a point  $x$  is simple if  $N_n(x) = N_p(x) = 1$ , where  $N_n(x)$  is the number of connected regions in the neighborhood of  $x$  with intensities less than or equal to the intensity at  $x$ . Likewise,  $N_p(x)$  denotes the number of connected regions in the neighborhood of  $x$  with intensities strictly greater than the intensity at  $x$ . The topology correction methodology used in this work successively detects and removes non-simple points for all isovalues of the input scalar field.

The process begins with an initialization with the desired topology which is propagated through successive isovalues using a fast-marching approach, the details of which can be found in section 2.2 of [2]. The points in the boundary around the initialized object are sorted by intensity and organized in a binary tree. At the  $k^{\text{th}}$  iteration, the first point is removed from the tree and added to the object if it is simple.

If the point  $x$  has been labeled as non-simple, its value is set to  $g(x) = \min_{y \in G_{k-1} \cap N(x)} g(y)$  where  $N(x)$  denotes the neighbors of  $x$  and  $G_{k-1}$  denotes the object at the  $k - 1^{th}$  iteration.

The correction can be done in two ways, “upward” or “downward.” The former is initialized in regions of high probability and raises areas of low probability, whereas the latter flattens regions of high probability and is initialized in regions of low probability. Figure 1 shows a simple 1D example of topology correction on scalar functions.

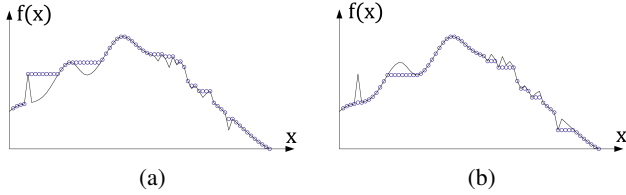


Figure 1. An example of topology correction applied to a 1D scalar function. Displayed in black are the original scalar functions and in blue, the corrected function for both a) upward correction and b) downward correction

Our method applies both topology correction schemes and chooses the result for which the sum of squared differences across all voxels is minimized. Formally, if  $\mathbf{V}_u$  and  $\mathbf{V}_d$  denote the “upward” and “downward” topology corrected probability maps, we choose our final probability map  $\mathbf{V} = V_i$  using:

$$\mathbf{V} = \arg \min_{Y \in \{\mathbf{V}_u, \mathbf{V}_d\}} \sum_i (W_i - Y_i)^2. \quad (3)$$

In this way, the resulting probability map is the closest projection (with respect to Euclidean distance) of the original on the space of topologically correct probability maps. An experiment is shown in section 3.1 that highlights the importance of approaching the problem in this way.

### 2.3. M-Step

Finally, the rater performance parameter  $\Theta$  that maximizes the conditional expectation of the complete data log-likelihood is computed. The details of the derivation can again be found in [18] or [14], and result in the following update equation

$$\theta_{rs's}^{(k+1)} = \frac{\sum_{i: D_{ir}=s'} W_{si}^{(k)}}{\sum_i W_{si}^{(k)}}. \quad (4)$$

Our use of the topology corrected probability map results in higher performance parameter estimates for raters that produced topologically correct objects and lower estimates for those whose labels were topologically defective. This

translates into a higher weighting for the raters with correct topology in the subsequent E-Step, resulting in a new probability map that is closer to the correct topology.

### 2.4. Overview

To summarize, Topology STAPLE carries out expectation-maximization (EM) iterations with

1. E-Step - estimate probability of object at each voxel
2. Project the resulting probability map onto the subspace of topologically correct probabilities
3. M-Step - estimate rater performance given the topologically correct probability map
4. Check for convergence

Convergence is detected by measuring the normalized trace of  $\Theta$  between subsequent iterations. The normalized trace averages the true positive and true negative fractions across raters and is computed by  $\frac{1}{2R} \sum_{r=1}^R \text{tr}(\theta_r)$ . Since the topology corrected probability maps and hard segmentations are of interest in this study, the E-Step and topology correction were performed once more after the final rater performances were computed.

### 2.5. Measuring Topology

The topological characteristics of the raters and the rater fusion results are described using the number of object parts, the number of object cavities, the number of object handles, and the Euler characteristic. The number of object parts ( $P$ ) is the number of connected components of the object. The number of object cavities ( $C$ ) is the number of connected components of the background completely enclosed within the object. A torus or coffee mug are examples of objects with one object handle ( $H$ ). The Euler characteristic is defined by  $\chi = V - E + F$  where  $V$  is the number of vertices,  $E$  is the number of edges, and  $F$  is the number of faces of polygonization of the object [7]. It is also related to the number of object parts, cavities and handles by:  $\chi = 2P + 2C - 2H$ .

In most of the studies described here, the true object is presumed to have spherical topology (that is,  $P = 1$ ,  $C = 0$ ,  $H = 0$ , and  $\chi = 2$ ). Three types of topological “defects” can occur in three dimensions. The object may have multiple disjoint pieces, resulting in  $P > 1$ . The object may contain cavities, resulting in  $C > 0$ . Finally, the object may possess “handles,” resulting in  $H > 0$ . The Euler characteristic is reported along with these three quantities whenever the topology of objects is reported.

## 3. Experiments

This proposed method was applied to both simulated and real data to gauge its efficacy and effectiveness. For

all experiments, the reported topology measures used 6-connectivity for the object, and 18-connectivity for the background, though similar results were obtained using other connectivity rules. The STAPLE algorithm was initialized with  $\theta_{rjj} = 0.9999 \forall r, j$ . The algorithm was determined to have converged when the difference of the normalized trace (see Section 2.4) between iterations was less than  $1 \times 10^{-5}$ .

### 3.1. Simulations

Topology STAPLE and STAPLE were run on simple examples to demonstrate the efficacy of the topology correction methodology. A shape resembling a gray matter gyral crown and sulcal fundus was created to mimic a situation in which errors in topology are likely to occur. Five synthetic raters were created by perturbing the boundary of the shape. Figure 2 shows (a) the true segmentation, (b) one example rater, (c) the STAPLE hard segmentation, and (d) the Topology STAPLE hard segmentation. From these results, it is evident that our method does indeed enforce the correct topology on the resulting hard segmentation.

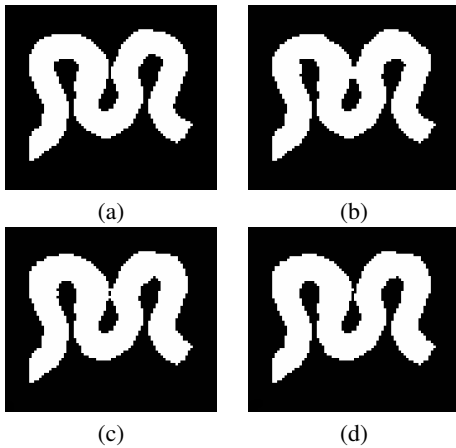


Figure 2. Simulation experiment - sphere/gyrus a) True label configuration b) Sample Rater c) STAPLE hard segmentation d) Topology STAPLE hard segmentation

These results demonstrate the ability of Topology STAPLE to produce topologically correct segmentations in the presence of rater error. Notice that object was “cut” rather than the handle filled. This is because our algorithm chooses the topology correction method that results in a probability map “closer” to the computed conditional probabilities, as described in section 2.2.

Finally, we note that it is possible to recover non-spherical topologies using our method, and demonstrate this ability with another simple example. This synthetic example resembles the white matter of the spinal cord and is topologically equivalent to a torus, or doughnut. The key to obtaining non-spherical topologies is to initialize the topol-

ogy correction step with a mask of the desired topology. Figure 3(a) shows the truth model with toroidal topology, (b), a sample rater in white and the initialization in red. Note that the initialization is of the correct topology and is “close” to the true segmentation, in the sense that the cavity of the initialization encloses the cavity in the truth model.

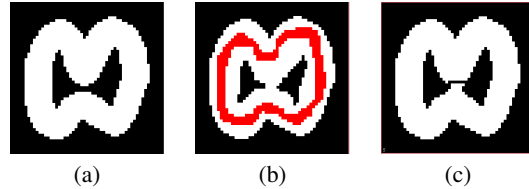


Figure 3. Simulation experiment - tous/spine a) True label configuration with a toroidal topology b) Sample Rater (white) and initialization (red) d) Topology STAPLE hard segmentation

Again, the segmentation produced from Topology STAPLE has the correct topology, whereas most raters and therefore, the STAPLE segmentations are topologically defective.

### 3.2. Real Data

Both Topology STAPLE and STAPLE were applied to real labelings produced by human raters. Four raters labeled the human cerebellum for 22 subjects from MP-RAGE MRI scans. The subject pool consisted of both control subjects and patients diagnosed with cerebellar degeneration. These four delineations for each subject were inputs to both STAPLE and Topology STAPLE. Table 1 shows several topological properties for all human raters, as well as for the resulting STAPLE and Topology STAPLE segmentations. The measures of topology described in section 2.5 were computed for each of the rater segmentations and for the STAPLE and Topology STAPLE segmentations. We note that none of the input raters have spherical topology as evidenced by the measures shown below. While in some cases, the STAPLE segmentation improves upon the rater results, as evidenced by fewer object parts and cavities, many topological defects remain. Topology STAPLE produces a segmentation with the correct topology for all cases.

Figure 4 shows a coronal slice of the segmentation one subject from all four raters, STAPLE, and Topology STAPLE. It is evident the segmentation from all four raters contained a topological error in the form of a cavity in the middle of the cerebellum. As a result, the labeling produced from the maximum a posteriori label probability obtained from STAPLE contains this hole as well. By enforcing our prior knowledge that the cerebellum has spherical topology, Topology STAPLE produces a segmentation estimate free of any cavities or handles.

Table 1. Topological Characteristics of Raters, STAPLE (ST) results, and Topology STAPLE (TST) results

	Sub.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
Euler Char.	1	44	100	40	36	76	108	142	106	150	96	48	166	66	46	18	104	20	66	96	64	104	272
	2	14	18	-38	48	22	30	42	-4	40	16	28	126	-4	32	22	28	-48	68	14	22	12	-22
	3	14	-8	42	76	50	26	98	52	114	52	26	120	22	42	20	60	-38	50	28	44	66	98
	4	28	42	8	40	26	96	66	46	62	30	50	136	10	66	6	32	-34	28	-2	18	6	6
	ST	10	12	-24	60	26	76	40	44	42	36	46	112	2	38	6	40	-18	40	6	26	12	56
TST	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Object Parts	1	48	78	99	55	56	47	82	58	63	71	42	52	81	44	10	51	44	46	90	46	37	178
	2	3	22	8	6	7	3	10	5	7	5	4	7	5	9	1	3	3	2	3	2	4	7
	3	10	24	64	40	28	23	65	26	22	22	14	18	21	35	8	31	35	10	38	34	42	74
	4	8	42	31	10	9	21	22	16	18	20	14	23	32	20	1	6	13	4	6	7	1	38
	ST	7	29	15	12	9	17	14	18	10	22	14	9	13	15	1	9	12	8	18	9	7	58
TST	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Object Cavities	1	16	33	47	52	17	50	38	41	81	46	24	102	31	33	6	48	28	31	34	18	39	37
	2	12	35	53	53	17	21	35	34	56	30	27	73	36	24	11	32	42	55	59	30	17	23
	3	17	43	64	55	20	30	51	42	83	47	27	71	46	33	5	49	18	30	29	26	18	37
	4	27	35	57	57	18	45	32	46	48	22	29	65	35	25	4	32	34	30	44	21	14	36
	ST	17	36	53	54	20	42	34	44	47	21	27	71	37	31	4	32	35	28	36	27	16	37
TST	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Object Handles	1	42	61	126	89	35	43	49	46	69	69	42	71	79	54	7	47	62	44	76	32	24	79
	2	8	48	80	35	13	9	24	41	43	27	17	17	43	17	1	21	69	23	55	21	15	41
	3	20	71	107	57	23	40	67	42	48	43	28	29	56	47	3	50	72	15	53	38	27	62
	4	21	56	84	47	14	18	21	39	35	27	18	20	62	12	2	22	64	20	51	19	12	71
	ST	19	59	80	36	16	21	28	40	36	25	18	24	49	27	2	21	56	16	51	23	17	67
TST	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

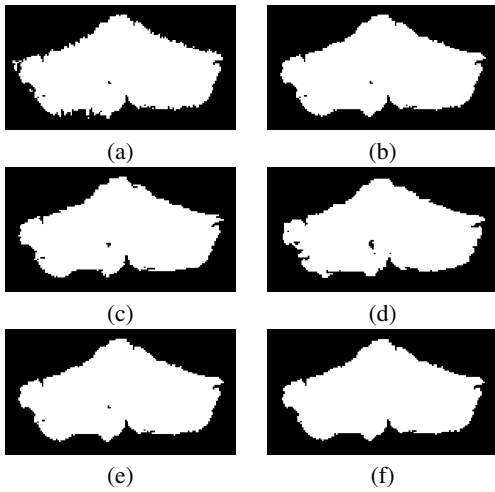


Figure 4. Experiment with real rater delineations of the cerebellum a) Rater 1, (b) Rater 2, (c) Rater 3, (d) Rater 4 delineations, (e) STAPLE hard segmentation (f) Topology STAPLE hard segmentation. The topology-corrected segmentation filled a number of cavities and handles that were ignored by STAPLE.

### 3.3. Multi Label Data

Finally, an experiment was carried out for which four raters annotated a volume with four object labels plus the background label. The objects under consideration were the anterior lobe, middle lobe, caudal lobe, and the corpus medullare of the cerebellum. We seek an estimate of the true label configuration for which all individual objects have spherical topology. This was done using a “confusion matrix” parameterization which generalizes the single-object, true-positive, true-negative parameterization. Like-

wise, equations 2 and 4 were generalized to multiple objects as shown in [14]. Topology correction was done individually on the probability maps for each object. Figure 5 (a) shows a surface rendering of volumetric labels obtained using Topology STAPLE. Figure 5 (b) and (c) shows a close up of part of the surface obtained from STAPLE and Topology STAPLE results respectively. A handle is visible in the former, but not in the latter.

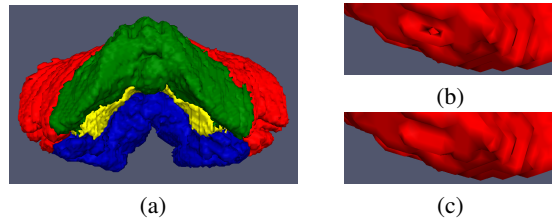


Figure 5. Multiple-label Topology STAPLE experiment. a) Topology STAPLE result using four objects, (b) A close up of a handle in one object of a STAPLE segmentation, (c) The same portion of the surface obtained from the Topology STAPLE segmentation.

We emphasize that by topology correcting the probability maps of each object individually, the resulting scalar maps are no longer memberships or probabilities, as the sum of the memberships are in general not equal to one. Furthermore, a topologically correct hard segmentation is not guaranteed when it is obtained by a max-membership rule, even if all of the memberships are topologically correct. Therefore, though closer to the correct topology than any individual rater or a STAPLE segmentation, the objects obtained using the above strategy will not have spherical topology in general, a challenge that will be addressed in future work.

## 4. Conclusion

This paper presents a novel method, Topology STAPLE, for estimating the hidden, topologically correct, true segmentation of an object and rater performances given rater delineations. Topology STAPLE was demonstrated to produce a topologically correct label estimate on synthetic data sets as well as on real cerebellar delineations from human raters. We have shown that none of the human raters, nor the STAPLE estimate applied on these results produced the correct topology. We verified on these data that our method produce a specified topology in practice. Our method was also applied on a dataset with multiple object labels, by applying topology correction to each label individually and updating probabilities accordingly.

Future work will study estimating the global, closest projection onto the topologically correct space. This could be done using graph-based correction methods [9, 16], for example, or by further development of the current fast marching correction method. Past work in the topology of multiple objects has employed “topology templates,” or initializations of the correct topology that are deformed according to membership functions. We will investigate the use of such templates and alternative techniques into a multi-object, topology-preserving version of STAPLE. With this tool we will be able to guarantee topological equivalence between segmentations obtained by combining multiple atlas-based labelings.

## References

- [1] J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston. Identifying global anatomical differences: deformation-based morphometry. *Hum. Brain Mapp.*, 6(5-6):348–357, 1998.
- [2] P.-I. Bazin and D. L. Pham. Topology Correction of Segmented Medical Images using a Fast Marching Algorithm. *Comput Methods Programs Biomed*, 88(2):182–190, 2007.
- [3] P.-L. Bazin and D. L. Pham. Topology-preserving tissue classification of magnetic resonance brain images. *IEEE transactions on medical imaging*, 26(4):487–96, April 2007.
- [4] G. Bertrand. Simple points, topological numbers and geodesic neighborhood in cubic grids. *Pattern Recognition Letters*, 15(10):1003–1011, 1994.
- [5] M. Couprie, F. N. Bezerra, and G. Bertrand. Topological operators for grayscale image processing. *Journal of Electronic Imaging*, 10(4):1003, 2001.
- [6] P. Golland, W. E. L. Grimson, and R. Kikinis. Statistical Shape Analysis Using Fixed Topology Skeletons: Corpus Callosum Study. *In Proc. IPMI, LNCS*, 1613:382–387, 1999.
- [7] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley, 1992.
- [8] U. Grenader and M. I. Miller. Computational anatomy: an emerging discipline. *Quarterly of Applied Mathematics*, 56(4):617–694, 1998.
- [9] X. Han, C. Xu, U. Braga-neto, and J. L. Prince. Topology Correction in Brain Cortex Segmentation Using a Multiscale, Graph-Based Algorithm. *IEEE Trans. Med. Imag*, 21(2):109–121, 2002.
- [10] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23 Suppl 1:S151–60, 2004.
- [11] T. Langerak, U. van Der Heide, I. Lips, A. Kotte, M. van Vulpen, and J. Pluim. Label fusion using performance estimation with iterative label selection. *2009 IEEE ISBI*, pages 1186–1189, June 2009.
- [12] J.-F. Mangin, V. Frouin, I. Bloch, J. Regis, and J. Lopez-Krahe. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *Journal of Mathematical Imaging and Vision*, 5(4):297–318, 1995.
- [13] D. Mitchen and C. Gaser. Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution. *Frontiers in neuroinformatics*, 3(25), 2009.
- [14] T. Rohlfing, D. B. Russakoff, and C. R. Maurer. Performance-Based Classifier Combination in Atlas-Based Image Segmentation Using Expectation-Maximization Parameter Estimation. *IEEE Trans. Med. Imaging*, 23(8):983–994, 2004.
- [15] A. I. Scher, Y. Xu, E. S. C. Korf, L. R. White, P. Scheltens, A. W. Toga, P. M. Thompson, S. W. Hartley, M. P. Witter, D. J. Valentino, and L. J. Launer. Hippocampal shape analysis in Alzheimer’s disease: a population-based study. *NeuroImage*, 36(1):8–18, 2007.
- [16] D. W. Shattuck and R. M. Leahy. Automated graph-based analysis and correction of cortical volume topology. *IEEE Trans. Med. Imag*, 20(11):1167–1177, 2001.
- [17] P. M. Thompson and A. W. Toga. A framework for computational anatomy. *Computing and Visualization in Science*, 5(1):13–34, July 2002.
- [18] S. K. Warfield, K. H. Zou, and W. M. Wells. Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation. *IEEE Trans. Med. Imaging*, 23(7):903–921, 2004.