The Impact of Atlas Formation Methods on Atlas-Guided Brain Segmentation

Lilla Zöllei[1], Martha Shenton[2], William Wells[2], and Kilian Pohl[2]

[1] Martinos Center, MGH, Boston, MA, USA; lzollei@nmr.mgh.harvard.edu
[2] Radiology Department, BWH, USA; shenton|sw|pohl@bwh.harvard.edu

Abstract. We analyze the impact of atlas construction within the context of an atlas-guided segmenter applied to a morphometry study in neuroanatomy. Automatic segmenters often rely on anatomical information encoded via probabilistic atlases. These atlases are frequently constructed by registering collections of training data. In this paper, we study the impact of registration methods as well as the training data on automatic segmentation results. With respect to registration, we focus our comparison on pairwise vs. group-wise methods and fixed vs. online coordinate systems. For the training data, we consider collections of population specific and general population data. To study the impact of these factors, we revisit a previously published statistical group comparison that was based on manual segmentations. For each atlas type, we record the group differences based on automatic segmentations and compare these findings to the original ones. Furthermore, we measure the Dice overlap between manual and automatic segmentations. Our results indicate some advantages for coordinate systems that are developed in an online fashion.

1 Introduction

Neuroscience studies frequently use volumetric measures of brain structures for the detection of morphological differences between patient groups [1]. These models are often based on manual segmentations, where experts outline substructures of major tissue compartments in MR images. Recently, automatic methods [2–5] have been proposed to replace this time-consuming procedure. These methods typically rely on statistical atlases (or spatial priors) to represent the variations within a population and to compensate for missing structural information in MR images.

Registration algorithms that facilitate the construction of these atlases can be characterized by, for example, their scalability, unbiasedness [6–8] and inverse consistency [9]. Although it is a difficult task to quantitatively characterize the performance of these methods with respect to one another, one can study their impact on particular applications. In this article, we specifically investigate how atlas creation procedures influence brain segmentation results.

Atlases used by segmentation processes are not only influenced by the type of registration but also by the nature of the underlying training data. For instance in a population-specific study, one might separately construct an atlas for each population, such as healthy and diseased, to increase the accuracy of the corresponding automatic segmentations. However, this might introduce a bias in the study as the data is not uniformly processed. On the other hand, if one chooses the conservative approach of building a single atlas based on a subset of the data, one risks under-representing the variations within the population [10]. Therefore it is essential to understand how atlas construction influences the segmentations. One recent study evaluates the performance of atlas-based segmentation methods using single (individual or average) vs. multiple

atlases [11]. In that case, however, the segmentation and atlas generation steps are combined, thus it is not possible to test for individual performance of these components.

In this paper, we study the impact of atlases by revisiting the study of [12], which analyzes morphometric differences in sub-structures of the gray matter between controls, first-episode affective, and first-episode schizophrenics. Instead of manual segmentations, we base the group comparisons on the results of an automatic segmentation tool [4] available through the 3DSlicer (http://www.slicer.org). More specifically, this tool incorporates statistical atlases capturing the spatial distribution of the structures to be segmented. By comparing the automatically generated labels to the gold standard manual segmentations of [12], we are able to quantify the influence of atlases on the outcome. We investigate three different methods for atlas construction using fixed vs. online coordinate systems and group- vs. pair-wise registration methods. Using those algorithms, both group-specific (multiple) and single atlas segmentations are carried out. To our knowledge, this is the first time that the impact of these various styles of atlas creation algorithms is quantitatively compared in an application. We compare our results to those that rely on the manual segmentations.

2 Atlas Construction

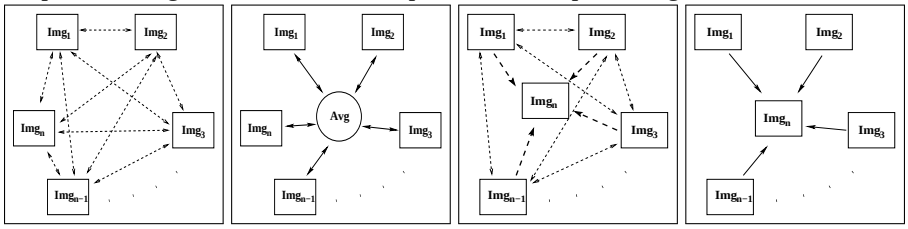
The goal of this paper is to study the impact of different atlas construction styles on the segmentation approach of [4] by revisiting the clinical study of [12].

As is done frequently, we first register the MR images of the training data set to a common coordinate frame. Then, the resulting transformations are applied to the corresponding manual segmentations. Label-specific probabilistic atlases are then defined by the normalized frequency of the occurrence of the label at each voxel location within the set of aligned segmentations. Finally, we align the atlas to the test cases by first registering the common coordinate system of the atlas to the MR images of the test cases and then applying the same deformations to the frequency maps. For this task, we choose the B-spline implementation by [13] (other intensity based - non-rigid registration methods would have been suitable choices). The brain segmentation tool then relies on the so-aligned frequency maps in order to resolve any potential ambiguities.

In this analysis, we differentiate between two sets of registration methods. The coordinate system of one set is fixed while it is developed online, during the execution of the algorithm, for the other. Some representatives of the former group are algorithms that use the Talairach coordinates frame [14, 15], and group-wise registration methods where a particular sample of the input population is selected in order to serve as a reference frame [16, 17]. In these methods, selecting the fixed template as a good group representative is non-trivial. For example, in the case of clinical studies, diseased members of the population could be erroneously chosen to represent the full data set. The latter registration group is represented by methods that are, by construction, unbiased. In their case the target coordinate system is either implicitly or explicitly computed during the alignment process [7, 6, 8, 18]. There exist methods that could be categorized in-between these two groups. Certain approaches, for example, first select a fixed registration reference frame and then, following a set of pair-wise alignments, the average transformation is applied to the reference in order to construct the final atlas [9, 19]. And in [20] the template is only used as an intensity reference and not for shape.

In the current study, we selected representatives of both groups in order to test for differences in the segmentation results due to the corresponding atlases. We implemented two registration methods that rely on a fixed coordinate system. One does a simultaneous group-wise registration to a particular member of the training set [16] and the other carries out a sequence of pair-wise registrations to a pre-selected template. For online alignment, we implemented an approach in which the mean of the input images is considered as the template and a set of pair-wise registrations ensure the group-wise spatial alignment (this technique is similar to the framework of [21]). The other online method is a simultaneous, unbiased group-wise registration technique referred to as congealing [8]. When selecting these techniques, we considered scalability, complexity, and computational performance of these algorithms. The pair-wise registration methods use mutual information [22] as an objective criterion and the other methods optimize the sum of voxel-wise entropies. The transformations that were recovered by the registration algorithms were 9-parameter affine. A brief summary and a schematic representation (see Fig.1) of our atlas construction methods follow:

- **GroupOnline** : do intensity-based group-wise registration among all the inputs simultaneously using sum of voxel-wise entropies [8]
- **PairOnline** : iterate between computing the mean of the input images (template) and performing intensity-based pair-wise registration between the template and the inputs based on mutual information
- **GroupFixed** : align training data to a fixed template through group-wise registration, which minimizes the sum of voxel-wise entropies on binary images [23]
- **PairFixed** : align the training data to a fixed template through a intensity-based pair-wise registration between template and the inputs using mutual information



(a) GroupOnline (b) PairOnline (c) GroupFixed (d) PairFixed

Fig. 1. A schematic description of the four registration method being compared for atlas construction. Solid arrows indicate explicit pair-wise registration while the dotted ones refer to an implicit alignment.

3 Experimental Setup

In order to evaluate the impact of the different atlas construction approaches on the segmenter, we revisit the study by Hirayasu et al. [12]. This brain morphometry study compares the relative volumes of the right and left superior temporal gyrus (rSTG/lSTG), amygdala (rAMY/lAMY), and hippocampus (rHIP/lHIP) between first-episode schizophrenia patients, first-episode affective psychosis patients, and healthy comparison subjects. The relative volume of an anatomical structure is the ratio between the volume measured from the manual segmentation of the structure and the intra-cranial cavity.

Hirayasu et al. tested the null-hypothesis that the relative volumes of the anatomical structures do not significantly differ between the three groups. The outcome of the corresponding ANOVA test appears in the second column of Table 1. In the remainder of this paper, we view this outcome as the baseline and ground truth for our comparisons.

The data set consists of 50 subjects: 16 SchiZophrenia patients (SZ), 17 AFFective patients (AFF), and 17 CONtrol subjects (CON). The MR acquisition protocol (1.5 Tesla, GE Medical Systems) included an SPGR ($256\times256\times124$, $0.9375\times0.9375\times1.5$mm) and and a T2-weighted sequence ($256\times256\times54$, $0.9375\times0.9375\times3$mm). Manual segmentations for the six substructures were available for all of these data sets.

Our analysis of the impact of an atlas on the results of the group difference study consists of three steps. In the first step, the atlas is constructed via the methods discussed in Sec. 2 from the training set. As mentioned earlier, the training set itself can bias the final results. That is because we assume that the training data properly represents the spatial variations within a population. We therefore alter the size and composition of the training data in the following ways:

- **GrpTr:** We construct an atlas tailored for each group by applying the methods of Sec. 2. This approach assumes that the spatial variations of the three groups in the full data set are best captured by a set of atlases and not a single one.
- **AllTr:** The training set consists of the entire data set. Unlike in the case of GrpTr, this method assumes that the variations within the full population can be adequately represented by a single probabilistic atlas.
- **ConTr:**The training set is the control group. This construction assumes that a restricted sub-group of the entire data captures the variations within the population.

Our tests are based on the *leave-one-out* framework which helps to reduce the risk of introducing further bias into the segmentation results.

In the second step of the work flow, the test data is segmented via the automatic method of [4] guided by the previously generated priors. We choose this approach as it is, to our knowledge, the only publicly available segmenter for cortical structures that can integrate atlases other than the ones provided by its implementation. We explicitly define the posterior and anterior boundaries across structures via the label maps of [12] as this boundary is defined by anatomical landmarks that were not provided in the training data set. An expert can quickly perform this task by identifying the first slice and the last slice showing the fornix along the border of the lateral ventricle.

In the third step of the work-flow we repeat the morphology study of [12] and perform a volume overlap analysis. A detailed description is presented in the next section.

4 Comparison and discussion

Visually, the atlases associated with the four registration methods do not produce significantly different results. Thus we discuss the quality of the atlases based on two different quantitative measures. In Sec. 4.1, we repeat the statistical analysis of [12] using the automatic segmentations and rate the corresponding atlases by the agreement of the results to the original findings. In Sec. 4.2, we measure volume overlap between the automatic and manual segmentations across the fifty cases using the Dice metric.

4.1 Relation Between Detecting Group Differences and Atlas Type

Table 1 summarizes the outcome of the study using the manual label maps (Manual) as well as the automatic segmentations guided by ten different atlases. Except for the PairFixed method, whose evaluation is computationally the most expensive, all registration methods have been tested with atlases constructed using the GrpTr, AllTr and ConTr methods. In the case of the PairFixed method we present results for the ConTr atlas.

Region	Manual	GroupOnline			PairOnline			GroupFixed			PairFixed
		GrpTr [1]	AllTr	ConTr	GrpTr[1]	AllTr	ConTr	GrpTr[2]	AllTr	ConTr	ConTr
lSTG	**0.021**	<**0.001**	0.054	0.089	**0.005**	**0.041**	n.s.	**0.015**	n.s.	n.s.	n.s.
rSTG	**n.s.**	**n.s.**	0.03	0.003	**n.s.**	**0.064**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**
lAMY	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	<0.001	**n.s.**	**n.s.**	**n.s.**
rAMY	**0.048**	**0.024**	**0.042**	0.073	**0.009**	**0.046**	**0.042**	<**0.001**	**0.039**	**0.040**	0.057
lHIP	**0.005**	<**0.001**	**0.001**	<**0.001**	<**0.001**	**0.004**	**0.001**	<**0.001**	**0.009**	**0.007**	**0.002**
rHIP	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	**n.s.**	0.024	**n.s.**	**n.s.**	**n.s.**

[1] No significant bias was measured within the set of atlases
[2] Atlas-related bias was detected

Table 1. The table contains the p-value of our One-Factor ANOVA tests based on manual as well as automatic segmentations. The automatic results were produced using ten different atlases. The statistical comparison of the automatic segmentations only agreed with the original findings for three atlases, all of them based on the Online technology.

For each structure, we performed a One-Factor ANOVA test based on the relative volume measures between the three groups of the input (AFF, CON, SZ). The corresponding p-values are recorded in the table. We consider p-values below 0.05 as significant and we abbreviate p-values above 0.1 as n.s. (non significant). Entries in bold indicate agreement with all the original findings of [12], which are also shown in Table 1 (Manual). We first note that none of the atlases based on a sub-set of the study data (ConTr) detects significant differences in the lSTG, which is not consistent with Manual. This suggests that seventeen subjects do not properly represent the variations of that structure in the data set.

We next discuss the performance of the atlases based on a fixed template (GroupFixed and PairFixed). In none of the six cases do the results agree completely with the original findings. However, the results of GroupFixed-GrpTr suggest significant group differences in five instead of three compartments (produced by the manual analysis). The lower p-values might be due to bias introduced when training an atlas for each group separately. To test this possibility, we segmented the controls with the three group-specific atlases (AFF, CON, SZ) and performed an ANOVA test for each structure. In this test, the results of GroupFixed-GrpTr indicate significant differences for every structure but for the lSTG. These results indicate that atlases based on a fixed template negatively impact the performance of the automatic segmenter as a fixed template increases the bias in the automatic segmenter.

The statistical analysis of the online methods (GroupOnline and PairOnline) reinforces the previous observation as the results of three out of the six experiments fully agree with the original findings. Furthermore, for the atlases based on GrpTr, in the case of both the GroupOnline and PairOnline frameworks, no significant bias was detected when segmenting the controls with the three group-specific atlases. This shows that the automatic segmenter can reproduce the original findings of the original study as long as one carefully chooses the method for atlas construction. In addition, it confirms the notion that templates generated from a group of scans are less likely to increase the bias in an atlas then templates based on a single scan. Finally, we note that it is only GrpTr that produces results for both online methods that are consistent with the original findings; more importantly, it does so with a lower p-value than originally reported. This

Reg	GroupOnline			PairOnline			GroupFixed			PairFixed
	GrpTr	AllTr	ConTr	GrpTr	AllTr	ConTr	GrpTr	AllTr	ConTr	ConTr
lSTG	71.1	72.3	72.0	72.5	73.4	73.2	69.9	71.3	70.2	69.1
rSTG	70.0	70.9	70.7	72.1	71.9	72.0	69.7	71.0	69.3	68.8
lAMY	85.8	86.2	87.0	86.5	86.7	87.3	85.5	85.7	86.7	85.7
rAMY	85.0	85.2	84.9	85.0	85.3	85.1	85.0	85.1	85.2	83.9
lHIP	80.3	81.5	80.8	81.7	82.0	81.5	81.8	80.4	80.8	77.4
rHIP	82.7	82.9	82.5	83.0	83.2	82.9	81.7	81.3	81.3	80.1
AVG	79.1	79.9	79.6	80.1	80.4	80.3	78.5	79.1	78.9	78.1

(a) Dice measure in percent over entire data set

Reg	Manual	GroupFixed
	Multiple	ConTr
lSTG	76.7	70.1
rSTG	76.2	71.4
lAMY	87.7	85.3
rAMY	87.3	83.0
lHIP	86.2	82.6
rHIP	85.7	82.8

(b) Reliability data set

Table 2. (a) Mean Dice score in percent of the automatic methods (b) Mean Dice score in percent of GroupFixed-ConTr with the ones produced by two experts (Manuals) over four cases

indicates that using a set of atlases better represents the spatial variations across the fifty subjects than just using a single one.

Finally, Table 1 seems to suggest that the PairOnline method produces more reliable results than GroupOnline as not only the findings for GrpTr but also AllTr agree with the original ones. However, we discovered that the slightly better performance of PairOnline does not indicate a lower bias compared to the GroupOnline framework but it is rather a result of a smoother atlas generated by PairOnline. We find that the automatic segmenter produces more accurate labelings using smoother atlases, which is indicated by the lower p-value in the significant findings in AllTr vs GrpTr. On the one hand, smoother atlases better capture the variability within our population as training is based on incomplete data. On the other hand, bias introduced by the non-rigid registration aligning the atlas to the scan of a test subject is reduced with smoother templates as also reported by [24]. With respect to our study, the degree of smoothness of the atlases is measured by the total entropy of the atlas, which are higher for the PairOnline than GroupOnline scheme. We therefore are not able to conclude that the atlases generated by PairOnline have less bias than those constructed by GroupOnline.

4.2 Volume Overlap between Automatic and Manual Segmentation

A popular metric used for measuring accuracy of automatic segmentations is the Dice measure with respect to manual labellings. Table 2 (a) summarizes the mean pair-wise Dice score in percent over the fifty image volumes and their structures.

We note that for each type of atlas training, PairOnline and GroupOnline achieve the highest average score for all six structures. In contrast to previous findings, the analysis of the Dice scores suggests that using a single atlas (AllTr) is favorable over a group specific one (GrpTr). For all the methods where the corresponding tests were run, GroupOnline, PairOnline and GroupFixed, the atlas computed by training on all inputs (AllTr) achieves a slightly higher average Dice score than GrpTr. We also note that the PairFixed approach receives the lowest average Dice score in the ConTr experiments. The accuracy of the segmentation method depends on how well the spatial priors of the atlas capture the variations within the population and on the accuracy of the non-rigid warp of the atlas to the target (the input to be segmented). The accuracy of the non-rigid registration is generally influenced by the resolution of the atlas, with lower resolution favored. Thus, the difference in performance between GroupOnline and PairOnline could potentially indicate that the accuracy of the spatial alignment pro-

duced by PairOnline is slightly inferior to that of GroupOnline. To better interpret the mean Dice scores, we also analyzed the four volumes from our data set that were selected for inter-rater reliability tests in [12]. These cases were segmented by the rater of our current study as well as two additional experts. As in the case of automatic segmentations, we computed the mean Dice score for the two additional experts by comparing their segmentations to the label maps of the rater of this study. The combined score of the two experts as well as the mean score of GroupFixed-ConTr for those four particular cases are presented in Table 2(b). For each structure, the results of our implementation received slightly lower scores than the manual ones. If we now interpret the Dice score of the experts as an indication for the ambiguity of the boundary location of a structure, then the discrepancy between the two types of segmentations is directly linked to this ambiguity. The discrepancy is highest for the lSTG (6.6%) and rSTG (4.8%), where the manual raters received the lowest Dice scores (76.7 % and 76.2 %). The discrepancy is lowest for the lAMY (2.4%), where the experts receive the highest score (87.7%).

5 Conclusion

We conclude from these experiments that for relatively large training sets, registration methods with a fixed template seem to bias the results more than those generating a template online. For our data set, we also showed that confining the training set to the control group or constructing it by random sub-sampling did not adequately represent the morphological variations within the data set. We therefore based one class of atlases on the entire data set via the leave-one-out style framework. The automatic segmentations produced by the atlas based on this training set in combination with the pairwise and groupwise online methods achieved the highest average Dice score over all other tested combinations. A less conservative approach is one where an atlas is generated for each group separately. One has to carefully apply this strategy as it can produce segmentation results leading to overconfident findings, such as in the case of the group-wise registration with a fixed template. However, we did not detect any group biases in the atlas when generated via the GroupOnline and PairOnline methods. These frameworks not only fully agreed with the original findings but also did so with higher confidence. We conclude that one has to carefully design an atlas for a given application.

References

1. E. Sullivan, J. Rose, and A. Pfefferbaum, "Effect of vision, touch and stance on cerebellar vermian-related sway and tremor: A quantitative physiological and MRI study," *Cerebral Cortex*, vol. 16, pp. 1077–86, 2006.
2. B. Fischl, A. van der Kouwe, C. Destrieux, E. Halgren, F. Segonne, D. Salat, E. Busa, L. Seidman, J. Goldstein, D. Kennedy, V. Caviness, N. Makris, B. Rosen, and A. Dale, "Automatically parcellating the human cerebral cortex," *Cerebral Cortex*, vol. 14, pp. 11–22, 2004.
3. D. Tosun, M. Rettmann, X. Han, X. Tao, C. Xu, S. Resnick, D. Pham, and J. Prince, "Cortical surface segmentation and mapping," *NI*, vol. 23, pp. 108–118, 2004.
4. K. Pohl, S. Bouix, R. Kikinis, and W. Grimson, "Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework," in *ISBI*, pp. 81–84, 2004.
5. R. A. Heckemann, J. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *NI*, vol. 33, no. 1, pp. 115–126, 2006.

6. S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, pp. S151–S160, September 2004.
7. C. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, and C. Taylor, "A unified information-theoretic approach to groupwise non-rigid registration and model building," in *IPMI*, vol. 3565 of *LNCS*, pp. 1–14, Springer, July 2005.
8. L. Zöllei, E. Learned-Miller, W. Grimson, and W. Wells, "Large efficient population registration of 3D data," in *Proceedings of ICCV05, CVBIA*, 2005.
9. G. E. Christensen, H. Johnson, and M. Vannier, "Synthesizing average 3d anatomical shapes," *Neuroimage*, vol. 32, pp. 146–158, August 2006.
10. D. Blezek and J. Miller, "Atlas stratification," in *MICCAI*, vol. 4190 of *LNCS*, pp. 712–719, 2006.
11. T. Rohlfing, R. Brandt, R. Menzel, and C. J. Maurer, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," *Neuroimage*, vol. 21, no. 4, pp. 1428–1442, 2004.
12. Y. Hirayasu, M. E. Shenton, D. Salisbury, C. Dickey, I. A. Fischer, P. Mazzoni, T. Kisler, H. Arakaki, J. S. Kwon, J. E. Anderson, D. Yurgelun-Todd, M. Tohen, and R. W. McCarley, "Lower left temporal lobe MRI volumes in patients with first-episode schizophrenia compared with psychotic patients with first-episode affective disorder and normal subjects," *The American Journal of Psychiatry*, vol. 155, no. 10, pp. 1384–1391, 1998.
13. T. Rohlfing and C. R. Maurer, Jr., "Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees," *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 1, pp. 16–25, 2003.
14. D. Collins, *3D Model-Based Segmentation of Individual Brain Structures from Magnetic Resonance Imaging Data*. PhD thesis, McGill University, Montreal, Canada, 1994.
15. D. Van Essen, H. Drury, S. Joshi, and M. Miller, "Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces," in *National Academy of Sciences*, vol. 95, pp. 788–795, 1998.
16. S. Warfield, J. Rexilius, P. Huppi, T. Inder, E. Miller, W. Wells, G. Zientara, F. Jolesz, and R. Kikinis, "A binary entropy measure to assess nonrigid registration algorithms," in *MICCAI*, LNCS, pp. 266–274, Springer, October 2001.
17. H. Park, P. Bland, A. Hero, and C. Meyer, "Least biased target selection in probabilistic atlas construction," in *MICCAI*, vol. 2 of *LNCS*, pp. 419–426, 2005.
18. N. Kovacevic, J. Chen, J. Sled, J. Henderson, and M. Henkelman, "Deformation based representation of groupwise average and variability," in *MICCAI*, LNCS 3216, pp. 615–622, Springer, 2004.
19. P. Aljabar, K. Bhatia, J. Hajnal, J. Boardman, L. Srinivasan, M. Rutherford, D. Dyet, A. D. Edwards, and D. Rueckert, "Analysis of growth in the developing brain using non-rigid registration.," in *ISBI*, pp. 201–204, 2006.
20. K. Bhatia, J. Hajnal, B. Puri, A. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *ISBI*, pp. 908–911, 2004.
21. P. Lorenzen, B. Davis, and G. Gerig, "Multi-class posterior atlas formation via unbiased kullback-leibler template estimation," in *MICCAI*, vol. 3216 of *LNCS*, pp. 95–102, 2004.
22. W. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *MIA*, vol. 1, pp. 35–52, 1996.
23. S. Warfield, J. Rexilius, P. Huppi, T. Inder, E. Miller, W. Wells, G. Zientara, F. Jolesz, and R. Kikinis, "A binary entropy measure to assess nonrigid registration algorithm," in *MICCAI*, vol. 2208 of *LNCS*, pp. 266–274, 2001.
24. D. Robinson and P. Milanfar, "Fundamental performance limits in image registration," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. pp. 1185–1199, 2004.