

Atlas Renormalization for Improved Brain MR Image Segmentation Across Scanner Platforms

Xiao Han and Bruce Fischl*

Abstract—Atlas-based approaches have demonstrated the ability to automatically identify detailed brain structures from 3-D magnetic resonance (MR) brain images. Unfortunately, the accuracy of this type of method often degrades when processing data acquired on a different scanner platform or pulse sequence than the data used for the atlas training. In this paper, we improve the performance of an atlas-based whole brain segmentation method by introducing an intensity renormalization procedure that automatically adjusts the prior atlas intensity model to new input data. Validation using manually labeled test datasets has shown that the new procedure improves the segmentation accuracy (as measured by the Dice coefficient) by 10% or more for several structures including hippocampus, amygdala, caudate, and pallidum. The results verify that this new procedure reduces the sensitivity of the whole brain segmentation method to changes in scanner platforms and improves its accuracy and robustness, which can thus facilitate multicenter or multisite neuroanatomical imaging studies.

Index Terms—Brain atlas, brain imaging, computational neuroanatomy, magnetic resonance imaging (MRI) segmentation.

I. INTRODUCTION

THE identification and delineation of brain structures from magnetic resonance imaging (MR) brain images is an important task and has many applications in neuroscience, such as the study of brain development, the mapping of functional activation onto brain anatomy, and the analysis of neuroanatomical variability among normal brains. Brain structure segmentation and morphometric analysis is also helpful in clinical diagnosis of neurodegenerative and psychiatric disorders, treatment evaluation, and surgical planning.

Advances in MR imaging technology have allowed for greater precision in the assessment of morphometric properties of brain structures *in vivo*; however, automated segmentation

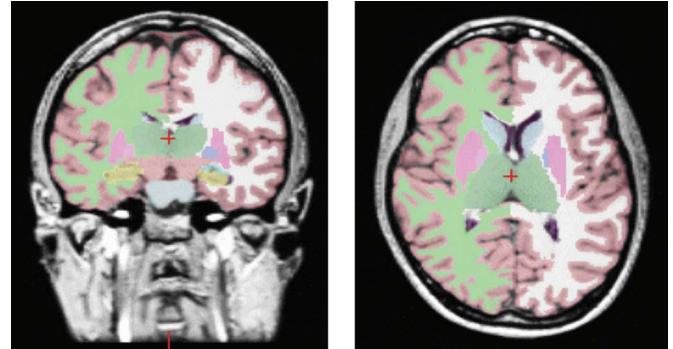


Fig. 1. Automated whole brain segmentation; label the brain into 37 structures.

and delineation of detailed brain structures remains a difficult task. The majority of existing efforts in MR brain image segmentation are devoted to the extraction and surface-based analysis of the brain cortex (cf., [1]–[7] and references therein), while work in the segmentation of noncortical structures is somewhat less common [8]–[13] and they are often tailored towards a small number of structures such as lateral ventricles, caudate, thalamus, and/or hippocampus.

In [12], Fischl *et al.* proposed an atlas-based segmentation method that automatically assigns labels to 37 subcortical structures (cf., Fig. 1), including left and right caudate, putamen, pallidum, thalamus, lateral ventricles, hippocampus, and amygdala. Unlike many other atlas-based approaches, the probabilistic atlas employed by the method maintains structure-specific intensity models at every atlas location in addition to the usual spatial prior for each structure. Compiling statistics separately for each subcortical structure at every atlas location prevents the broadening of underlying intensity distributions that would otherwise occur and thus allows for more accurate atlas registration and brain structure segmentation. The whole brain segmentation method was later extended to include a forward model of the MR image formation and to incorporate a nonlinear atlas registration scheme to improve accuracy in the atlas registration [14].

The reported method has been shown to provide segmentation accuracy comparable to manual labeling [12], [14]. Unfortunately, the performance often degrades when processing data acquired with a scanner platform and/or pulse sequence that differs from the training data used for the atlas construction, due to changes in the underlying image contrast and hence mismatch between the prior intensity model and the input image. This limitation hinders the application of the method in multicenter or multisite neuroanatomical studies that typically involve data acquired from multiple platforms. The same difficulty is faced by other atlas-based methods as well. The method of [14] partially

Manuscript received May 26, 2006; revised November 3, 2006. This work was supported in part by the National Center for Research Resources under Grants P41-RR14075 and R01 RR16594-01A1, in part by the NCRR BIRN Morphometric Project BIRN002, U24 RR021382, in part by the National Institute for Biomedical Imaging and Bioengineering under Grant R01 EB001550, and in part by the Mental Illness and Neuroscience Discovery (MIND) Institute, part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research under Grant U54 EB005149. Asterisk indicates corresponding author.

X. Han was with the Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129 USA. He is now with CMS, Inc., St. Louis, MO 63132 USA (e-mail: xiao.han@cmsrtp.com).

*B. Fischl is with the Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129 USA and also with the AI Lab, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: fischl@nmr.mgh.harvard.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2007.893282

addressed the problem by formulating the atlas model in the space of intrinsic tissue parameters, which, however, requires steady-state MR sequences and requires multispectral training data for tissue parameter estimation.

In this paper, we propose a procedure to automatically adjust the intensity model of an existing atlas to new image data, which results in improved accuracy and robustness of the overall method. This atlas renormalization procedure applies a multilinear atlas-image registration and a histogram matching step to update the class-conditional densities for each structure represented in the atlas (e.g., caudate, putamen, hippocampus, and etc.). The histogram matching step shares some similarity with existing methods on histogram-based MR image intensity normalization or standardization [15]–[19], but there are important differences between them in terms of the overall objective and methodology. First, the present work aims to develop a robust whole brain segmentation method and the normalization is performed on the model parameters of a given probabilistic brain atlas. On the other hand, the intensity standardization methods were proposed to normalize the intensity scales across multiple images, which mainly serve as a preprocessing technique that can facilitate visualization and subsequent tissue segmentation. Their application in detailed brain structure segmentation, however, remains to be investigated. Second, the intensity standardization methods adopted a global piecewise linear intensity mapping function estimated using global histogram features, which may not be sufficient to address nonlinear image contrast changes among individual structures across different imaging sequences. In this paper, we propose to correct the prior intensity models for each individual structure separately and a structure-wise atlas-image registration is further developed for better histogram estimation.

The work presented here also shares some common aspects with a class of brain segmentation methods based on the expectation-maximization framework (cf., [20], [21], and references therein) where the model parameters are iteratively updated during the tissue classification process. These methods typically focus on simultaneous bias field correction and tissue segmentation of gray matter, white matter, and CSF, and assume a spatially stationary intensity distribution for each tissue class. In general, this reduces segmentation accuracy as different parts of heterogeneous structures such as the thalamus and hippocampus have significantly different intensity distributions, resulting in intensity histograms across space that are not well approximated by Gaussians. In contrast, the framework of [12] and [14] allows the intensity model to vary for each individual structure and at each atlas location, allowing the modeling of internal structure such as the perforant path, a white matter bundle that traverses the hippocampus. Another set of related work is the recent development of the polyaffine registration framework by Arsigny *et al.* [22], which may provide more robust structure-wise linear registration than the multilinear approach designed in the current work, as it can guarantee invertibility of the composite transform.

In the following, we first summarize the whole brain segmentation method and then describe the intensity renormalization procedure. Finally, results are presented to demonstrate the improved robustness and accuracy of the new method. We note that

only T1-weighted image data and a single-channel atlas are used in the current work. Validation of the method on multichannel or multimodality brain segmentation (where the atlas intensity renormalization can be performed for each channel separately) is ongoing and the results will be reported in the future. The binaries for the method and the brain atlas used are available to the research community through the recent release of FreeSurfer at <http://www.surfer.nmr.mgh.harvard.edu/>.

II. ATLAS-BASED WHOLE BRAIN SEGMENTATION

In this section, we summarize the overall principle of the whole brain segmentation method proposed in [12] and [14]. The same framework is adopted in the current work, but with an additional step for automatically correcting the prior atlas intensity model to better fit new input data.

In [12] and [14], the problem of automatically labeling brain structures is phrased within the framework of Bayesian estimation theory, which allows for explicit incorporation of prior information about the spatial distribution of individual structures and their expected intensity appearance. In particular, both the priors on the anatomical labels (W) and the conditional probability $p(I|W)$ of observing the image given the classification are expressed within an atlas space, allowing them to vary as a function of position within the brain. In addition, the segmentation itself is modeled as a nonstationary anisotropic Markov random field (MRF). Due to the unknown transformation beforehand between the atlas space and the image to be segmented, the problem can be formally written as a maximization of the joint *a posteriori* probability of both the segmentation W and the atlas function f (i.e., the registration between the atlas and a given image)

$$\begin{aligned} (W, f) &= \arg \max_{W, f} p(W, f|I) \\ &= \arg \max_{W, f} p(I|W, f)p(W|f)p(f). \end{aligned} \quad (1)$$

In the above equation, the term $p(I|W, f)$ encodes the relationship between the class label at each voxel location over the image domain R and the predicted image intensity

$$\begin{aligned} p(I|W, f) &= \prod_{r \in R} p(I(r)|W(r) = c, f) \\ &= \prod_{r \in R} \frac{\exp\left(-\frac{1}{2} \|I(r)\mu_c(f(r))\|_{\Sigma_c^{-1}(f(r))}^2\right)}{(2\pi)^{p/2} \|\Sigma_c(f(r))\|^{1/2}} \end{aligned} \quad (2)$$

where $\mu_c(f(r))$ and $\Sigma_c(f(r))$ are the mean vector and covariance matrix, respectively, for the multichannel Gaussian intensity model at the atlas location $f(r)$ for class label c , and $\|I - \mu_c\|_{\Sigma_c^{-1}}^2 \equiv (I - \mu_c)^T \Sigma_c^{-1} (I - \mu_c)$. The term $p(W|f)$ is given by

$$\begin{aligned} p(W|f) &\propto \prod_{r \in R} p(W(r)|f(r)) \\ &\quad \times \prod_{r_i \in N(r)} p(W(r_i)|W(r), r_i, f(r)) \end{aligned} \quad (3)$$

which allows two types of prior information to be incorporated into the segmentation procedure. The approximate location a

neuroanatomical structure may occupy within the brain is given by $p(W(r)|f(r))$. The local relationship between anatomical classes is encoded in $p(W(r_i)|W(r), r_i, f(r))$, which denotes the pairwise probability that anatomical class $W(r_i)$ occurs at the r_i th neighbor of class $W(r)$. This probability depends not only on the neighborhood location r_i but also on the coordinates in the atlas space $f(r)$, thus gives a nonstationary anisotropic MRF model as opposed to isotropic ones typically used in the literature. These prior probabilities are learned during the atlas training stage, as detailed in [12].

The last term $p(f)$ provides a means for constraining the space of allowable atlas functions, which corresponds to the regularization typically employed in nonlinear atlas registration methods, e.g., continuity, differentiability, and invertibility. The explicit expression of $p(f)$ can be found in [14].

To reduce the computation complexity, the joint MAP estimation was divided into two major stages: computing the nonlinear atlas function and refining the segmentation. In the first stage, (1) is simplified by ignoring the MRF term, i.e., ignoring the neighborhood dependency of segmentation labels, which leads to a simpler iterative approach to estimate the atlas function. After the nonlinear atlas function is computed, an iterative conditional modes (ICM) algorithm is applied to refine the segmentation labels by incorporating the full MRF model but with the atlas function f being fixed.

III. ATLAS RENORMALIZATION METHOD

Because of the dependency of the image likelihood on the prior intensity model ($\mu_c(f(r))$ and $\Sigma_c(f(r))$) at each atlas location and for each anatomical class c , the accuracy of the atlas registration (computation of the atlas function f) and the final segmentation often degrades when the input image has different contrast properties than the training data used for the atlas construction. This happens when the scanner types or imaging sequences change during the process of a longitudinal neuroanatomical study or when multiple acquisition types are inevitably involved in a multicenter or multisite study. To address this problem, we propose to automatically adjust the prior intensity models of the atlas, in order to better fit the input data. The atlas model renormalization is performed for individual structures, and a linear intensity mapping is assumed between the input image and the atlas for each structure. The procedure is motivated by two underlying insights. The first is that while a global linear registration does not, in general, result in good alignment of structures in the atlas with their true location in the image, a structure-specific linear alignment does. The second is that while a Gaussian intensity model is a reasonably accurate one on a location by location basis, it does not provide a good description of the spatial distribution of intensities of a structure, which causes us to use nonparametric (histogram) techniques to infer the true distribution from a novel image. We will demonstrate later that these assumptions lead to significant improvement of the segmentation accuracy and result in a procedure that is as accurate for novel imaging data as it is for the data used to build the prior intensity models.

The atlas renormalization consists of a multilinear atlas-image registration and a histogram matching step. In the first step, a local linear registration is computed for each

individual structure, the result of which defines a (rough) delineation of the particular structure on the input image. The linear registration is defined as a 12-parameter affine transformation, which can be denoted in matrix form as $y = Mx$ or

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{pmatrix} = \begin{pmatrix} m_{1,1} & m_{1,2} & m_{1,3} & m_{1,4} \\ m_{2,1} & m_{2,2} & m_{2,3} & m_{2,4} \\ m_{3,1} & m_{3,2} & m_{3,3} & m_{3,4} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \quad (4)$$

where M denotes the affine transformation matrix.

In the second step, an intensity histogram is built from the image and the atlas separately for each structure, and a linear intensity mapping is computed to best align the two histograms. A linear intensity mapping is used because we assumed a Gaussian intensity model for each structure at each atlas location, and thus a linear mapping is sufficient to correct the mean (by the offset parameter of the linear mapping) and the standard deviation (by the scaling factor) of the prior class-conditional Gaussian intensity model for each structure. We detail the two steps in the following. Note that we assume single-channel input images and atlas in the following, although the extension to the multichannel case is straightforward.

A. Multilinear Atlas-Image Registration

The structure-wise linear (hence multilinear for the whole image) registration is initialized by a global affine registration that aligns the image to the atlas space. We note that an input MR image first undergoes some preprocessing stages before any of the registration is computed. The preprocessing isolates the brain from nonbrain tissues (also known as skull-stripping) and removes possible intensity bias field that often corrupts MR images. The preprocessing is done fully automatically as described in [3]. After the preprocessing, the white matter intensity is, in general, normalized around a fixed value. The same preprocessing was performed on all the training data used to construct the atlas as well.

The global linear registration finds the 12-parameter affine transform L that maximizes the likelihood of observing the image I given the atlas A , or equivalently, minimizing the residual of fitting the intensity model of the atlas to the data:

$$L = \arg \max_L \log p(I|A, L).$$

$p(I|A, L)$ has the same expression as in (2), where the label $W(r)$ is fixed to be the most probable class at the atlas location $L(r)$. The maximum likelihood optimization is solved using an iterative global search along each of the rotation, scale, and translation axes, followed by a Davidson–Fletcher–Powell (DFP) numerical maximization [23] using the gradient of $\log p(I|A, L)$. Due to the small degree of freedom (only 12), the global linear registration is insensitive to changes in image contrast.

The global registration is insufficient to accurately align each individual structure for the purpose of histogram matching. That is, the misalignment causes the histogram estimates of the densities to be significantly corrupted by other tissue classes. We thus further refine the registration for each structure individually. For this purpose, we first construct a template subvolume for each

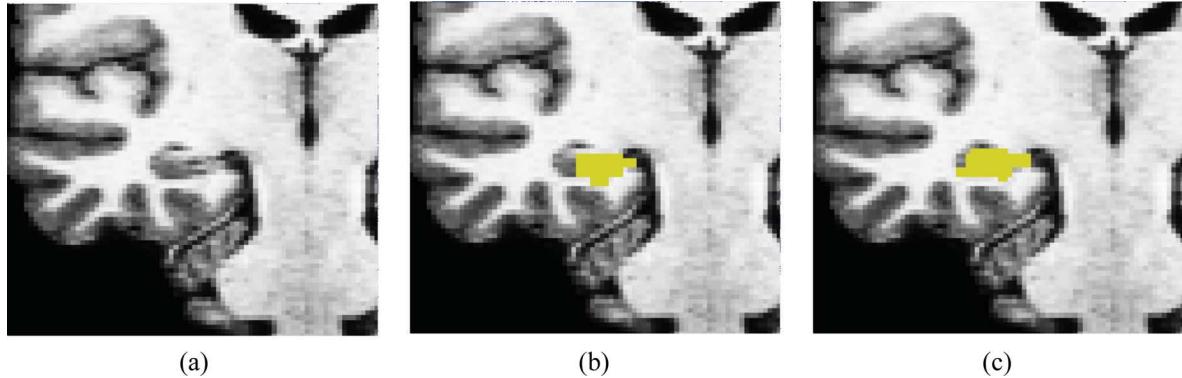


Fig. 2. Illustration of local linear registration for right hippocampus: (a) underlying image; (b) after global linear registration; and (c) after local refinement. Note that hippocampus region (yellow overlay) in (b) extends a few voxels into the white matter inferior to hippocampus, an error that is corrected in (c).

structure using the atlas. The template is simply a subregion of the atlas where the most probable class label is the same as the structure under consideration. To provide enough contextual information, we also dilate the initial template by a small radius (3 mm in our current implementation). Finally, the intensity value at each voxel of the template subvolume is assigned to be the mean intensity of the most probable class at that atlas location.

The template volume (T) is registered to the input image (I) by minimizing the following quadratic error function, which takes into account possible changes in image contrast and brightness:

$$E(\mathbf{M}) = \sum_i (I(\mathbf{M}\mathbf{x}_i) - cT(\mathbf{x}_i) - s)^2 \quad (5)$$

where i is the index of the voxels of the template volume. The scalars c and s are the parameters for contrast and brightness change, and \mathbf{M} denotes an affine transformation matrix as defined in (4).

Minimization of (5) involves generating a linear approximation to the sum-of-squared-difference function using Taylor's theorem [23] and solving the resulting matrix equation, which gives (cf., [24])

$$\mathbf{q} = (\mathbf{C}^T \mathbf{C})^{-1} (\mathbf{C}^T \mathbf{k})$$

where $\mathbf{q} = [m_{1,1}, m_{1,2}, \dots, m_{3,4}, c, s]^T$ is a vector for the 12 unknown affine transformation parameters and two unknown intensity factors. Element k_i of vector \mathbf{k} is equal to the value of $x_1 I_{x_1} + x_2 I_{x_2} + x_3 I_{x_3} - (I + T)/2$ evaluated at the i th voxel $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})$. The matrix \mathbf{C} is the Jacobian matrix of the residual $(I(\mathbf{M}\mathbf{x}) - aT(\mathbf{x}) - b)$, whose element $c_{i,j}$ is equal to the derivative of the i th residual (i.e., the residual at the voxel \mathbf{x}_i) with respect to the parameter q_j . In particular, the derivative of the i th residual with respect to element $m_{j,k}$ of matrix \mathbf{M} is equal to $(\mathbf{x}_i)_k (\partial I(\mathbf{y}) / \partial y_j)$ for elements $m_{1,1}$ to $m_{3,3}$ and simply $\partial I(\mathbf{y}) / \partial y_j$ for elements $m_{1,4}$ to $m_{3,4}$, where $(\mathbf{x}_i)_k$ denotes the k th component of the coordinate vector \mathbf{x} at voxel i . The derivative of the i th residual with respect to the scaling parameter c is simply $-T(\mathbf{x}_i)$, the negative intensity of the template volume T at the i th voxel. The derivative with respect to the offset parameter s is -1 for every voxel.

Since the Taylor approximation is most valid only when the estimation is close to the initial condition (i.e., the identity transformation), the estimated parameters need to be updated iteratively. For each iteration, the estimated transformation is applied to the template, and a new transformation is estimated between the newly warped template volume and the source image. The final transformation is obtained through a composition of the consecutive incremental transformations. As noted above, the estimation of the affine transformation for each individual structure starts with the same global registration as the initialization.

An example of the individual registration is shown in Fig. 2 for the right hippocampus, where the template volume is shown on top of the given image as an overlay. Note the inclusion of significant quantities of white matter voxels in the hippocampal region obtained from the global linear registration, a problem that is alleviated by the structure-specific alignment.

B. Histogram Matching

The registered template defines a mask for the corresponding structure in the input image, the intensity histogram of which is then built within the masked area for the particular structure. Another histogram for the same structure is also constructed from the atlas, using the mean intensity value at each atlas location for the given structure.

Next, we compute a linear intensity mapping that maximizes the correlation between the two histograms

$$(a, b) = \arg \max_{(a', b')} \sum_i h(i) \cdot g(a'i + b').$$

The optimal solution is computed through a global search of the parameter space. The search range is empirically set to $[0.3, 3]$ for a , and $[-255, 255]$ for b , and the step size is set to 0.1 for a and 0.5 for b . An example for the histogram matching is illustrated in Fig. 3, where we show the estimated histograms for the hippocampus from the image and the atlas, respectively, and the transformed atlas histogram after the linear matching. We note that although an intensity mapping was also estimated in the previous structure registration step, the computation there included all the voxels in the template subvolume (not only those for the particular structure), and thus the results were not as accurate.

To improve the robustness of the intensity renormalization, the local structure-wise linear registration and histogram

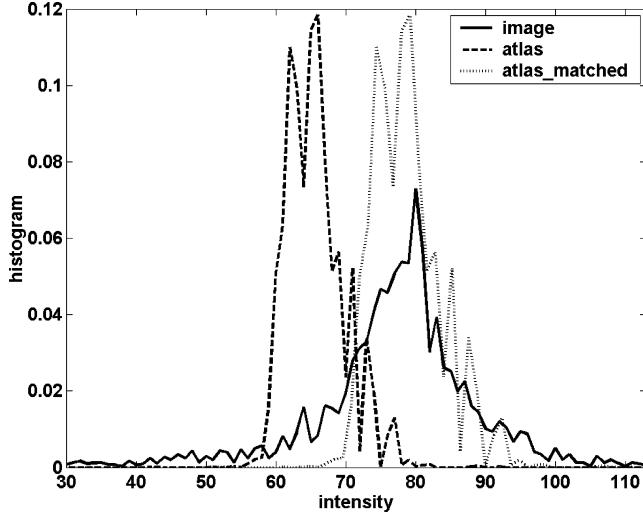


Fig. 3. Illustration of histogram matching. Plot shows histogram for hippocampus built from atlas and input image, respectively, and also atlas histogram after computed linear intensity mapping.

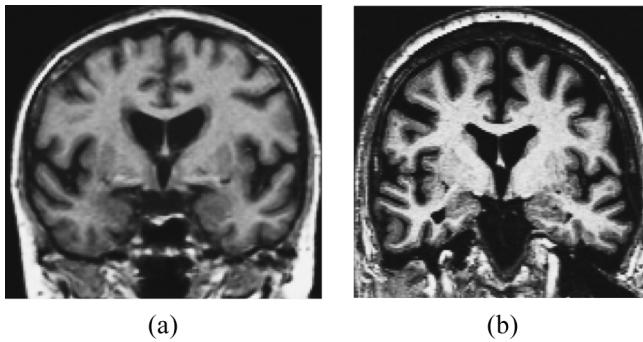


Fig. 4. Cross-sectional view of two different datasets: (a) GE dataset and (b) Siemens dataset.

matching are computed for major structures only including (left and right hemisphere computed separately) white matter (WM), lateral ventricles, thalamus, caudate, putamen, pallidum, hippocampus, amygdala, brain stem, third ventricle, and fourth ventricle. These structures are further grouped into three major tissue classes: WM, gray matter (GM), and CSF. As a result, we also estimate an average linear intensity mapping for each of the three tissue groups. The average mapping is assigned to other anatomical structures according to their tissue type. In addition, we automatically detect outliers in histogram matching results by rejecting the matching if the initial histogram overlap is small. In this case, the individual histogram matching for the particular structure is rejected, and the average mapping computed from other structures of the same tissue type is assigned to it in the end. After the histogram matching, the computed intensity transformation is applied to correct the mean and variance of the prior intensity model corresponding to the given structure across the whole atlas. In particular, the variance is scaled by a^2 , and the mean is transformed by $a\mu + b$. The corrected atlas intensity model is then applied to the subsequent nonlinear atlas registration and MAP labeling steps. In addition, the nonlinear registration is now initialized by the multilinear registration results instead

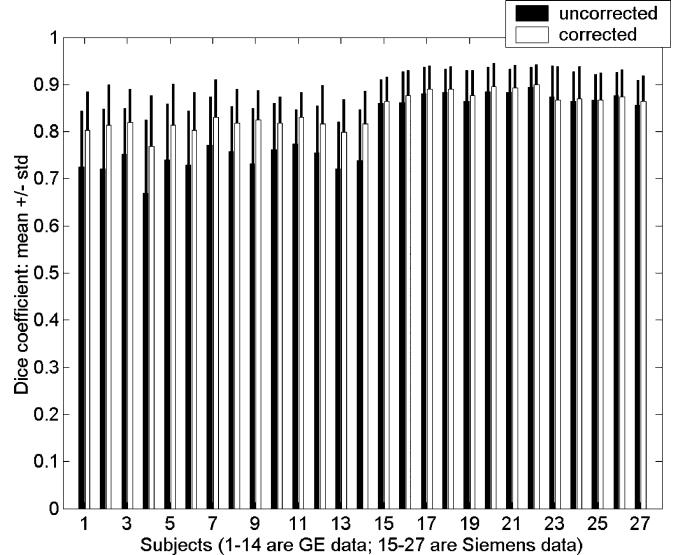


Fig. 5. Comparison of overall accuracy for 27 subjects. Bar plots show average Dice coefficients and error bars indicate one standard deviation.

of the global one, which we anticipate will result in a more robust and accurate procedure. Specifically, each atlas node is initialized with a movement vector that takes it towards the target location generated by applying the linear transform for the structure at that node to the node position. A series of line minimizations [23] is then performed using the standard energy functional in [14]. This results in an approximation to the full piecewise linear transform that is smooth and invertible; the subsequent refinement follows the same procedure as in the original method [14].

IV. RESULTS AND DISCUSSION

Two groups of T1-weighted MR brain images from a total of 27 subjects were used to test the proposed method. The first group consists of 14 data sets acquired on a 1.5T GE Signa scanner (General Electric, Milwaukee, WI) with the *SPoiled Gradient Recalled* (SPGR) sequence. The parameters are: TR = 35 ms, TE = 5 ms, flip angle = 45°, voxel size = $0.9375 \times 0.9375 \times 1.5 \text{ mm}^3$ (124 coronal slices). The second group consists of 13 data sets acquired on a 1.5T Siemens Sonata scanner (Siemens Medical Solutions, Erlangen, Germany) with the *Magnetisation Prepared, Rapid Acquisition Gradient-echo* (MPRAGE) sequence; and the parameters are: TR = 9.7 ms, TE = 4 ms, TI = 20 ms, flip angle = 10°, voxel size = $1.0 \times 1.0 \times 1.5 \text{ mm}^3$ (128 sagittal slices). Fig. 4 shows cross sections of one GE data set and one Siemens data set, illustrating the large differences in gray/white contrast between these acquisition types. We note that the atlas used in the experiments below was built from a separate set of 40 training data acquired on the same Siemens platform. Thus, the test data sets allow us to evaluate the performance of the proposed method both within and across scanner platforms and pulse sequences.

We applied the whole-brain segmentation method to label all the test data sets both with and without the atlas renormalization procedure outlined in this paper. To evaluate the segmentation

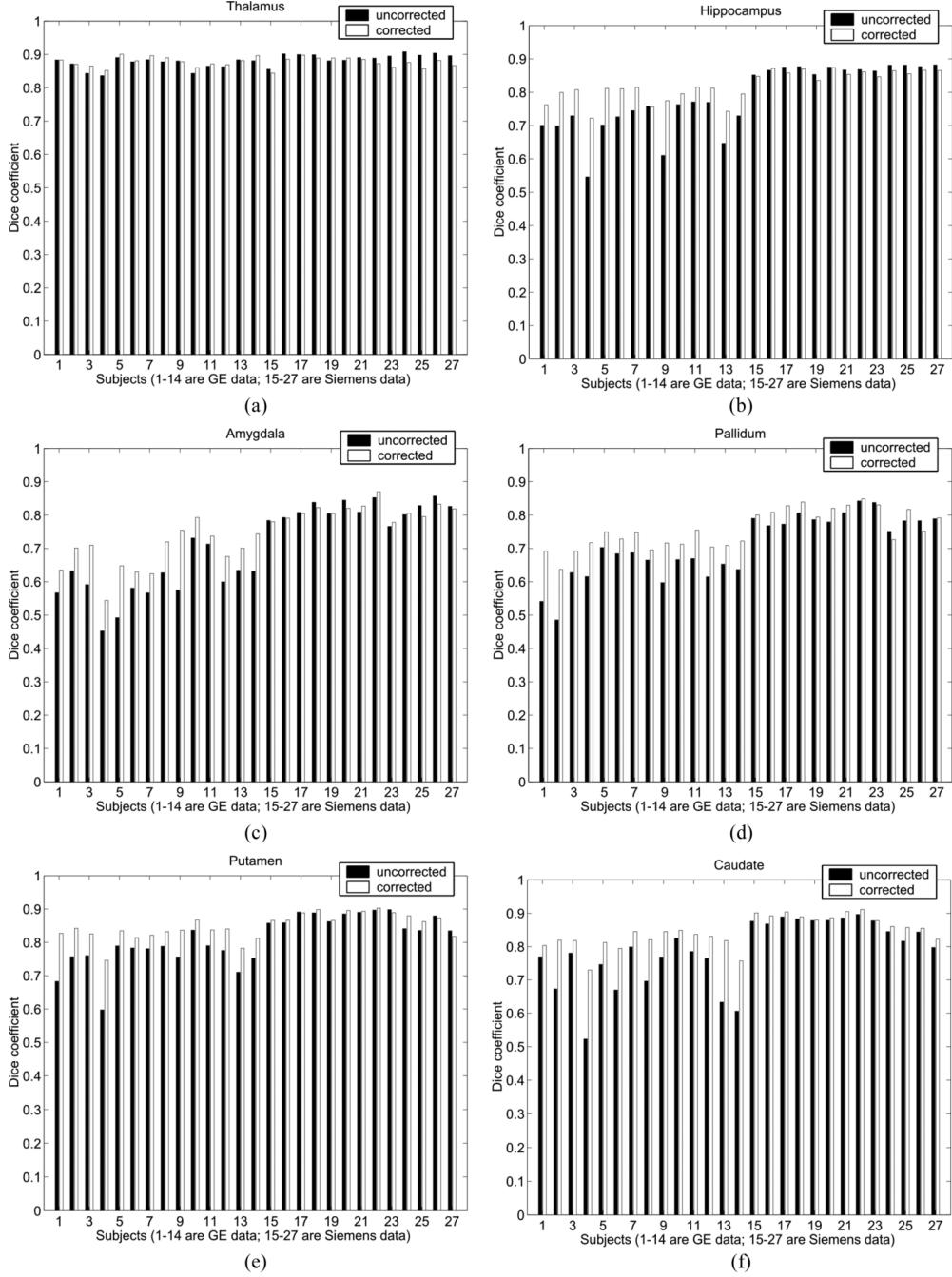


Fig. 6. Comparison of segmentation accuracy of two methods (with and without atlas intensity renormalization) for six individual structures: (a) thalamus, (b) hippocampus, (c) amygdala, (d) pallidum, (e) putamen, and (f) caudate.

accuracy, we compared the automated results with manual labeling available on the test datasets and computed the Dice coefficients [25] for the volume overlap. In particular, given two different labels (automatic and manual) of a structure, denoted by W_1 and W_2 , and a function $V(W)$, which takes a label and returns its volume, the Dice coefficient is given by [25]

$$D(W_1, W_2) = \frac{V(W_1 \cap W_2)}{(V(W_1) + V(W_2)) / 2}.$$

For identical labelings, $D(W_1, W_2)$ achieves its maximum value of one, with decreasing values indicating less perfect overlap.

The Dice coefficients were computed for all the structures. The overall mean and standard deviation of the Dice coefficients for the test subjects are summarized in Fig. 5, comparing the old and the augmented segmentation approaches. The Dice coefficients for six individual subcortical structures are also shown in Fig. 6, where an average was taken across the left- and right- hemisphere for the same structure. The group-wise mean and standard deviation of the Dice coefficients with and without the atlas intensity correction are also shown in Table I for the six individual structures and the average of all structures.

From the results we can see that the atlas renormalization procedure largely improves the accuracy of the automated

TABLE I
COMPARISON OF AVERAGE SEGMENTATION ACCURACY (DICE COEFFICIENTS) OF TWO GROUPS OF SUBJECTS WITH AND WITHOUT APPLYING ATLAS INTENSITY RENORMALIZATION. THE *p*-VALUES WERE COMPUTED USING STANDARD PAIRED-*t* TEST

		Thalamus	Hippocampus	Amygdala	Pallidum	Putamen	Caudate	All
GE	without correction	0.87±0.02	0.71±0.06	0.60±0.07	0.63±0.06	0.75±0.06	0.72±0.09	0.74±0.03
	with correction	0.88±0.01	0.79±0.03	0.69±0.06	0.71±0.03	0.82±0.03	0.81±0.03	0.81±0.02
	<i>p</i> -value	0.002	3e-5	3e-6	3e-6	1e-5	4e-5	
Siemens	without correction	0.89±0.01	0.87±0.01	0.82±0.03	0.79±0.03	0.87±0.02	0.86±0.03	0.87±0.01
	With correction	0.88±0.02	0.86±0.01	0.81±0.02	0.81±0.04	0.88±0.02	0.88±0.03	0.88±0.01
	<i>p</i> -value	0.003	0.004	0.3	0.07	0.15	0.004	

segmentation on the GE data sets, while the accuracy for the Siemens data sets remains at a high level (since the Siemens data came from the same scanner platform as the training data used to build the atlas). From the individual structure results, we can see that the sensitivity of the segmentation accuracy is structure dependent. The accuracy for thalamus is comparable across the GE and Siemens scanners and less affected by the atlas renormalization procedure. For the other five structures (hippocampus, amygdala, caudate, putamen, and pallidum), the segmentation accuracy can be quite poor for the GE data (average Dice coefficients between 0.60 and 0.75; cf., Table I) without applying the intensity renormalization procedure. The accuracy is improved by 10% or more (average Dice coefficients between 0.69 and 0.82; cf., Table I) with the newly developed intensity correction procedure incorporated into the registration and segmentation. As indicated by the *p*-values shown in Table I, the bulk of improvements for the GE data sets are statistically significant at a significance level below 0.01. For some structures, including thalamus, hippocampus, and amygdala, the segmentation accuracy is slightly degraded for the Siemens data sets when the atlas renormalization procedure is applied. This is because the renormalization procedure attempts to modify an originally correct (as reflected by the high initial Dice coefficients) atlas intensity model in such cases and can therefore introduce inaccuracy to the segmentation process. Although the degradation also reaches a statistical significance level of 0.01 for thalamus and hippocampus, its magnitude is minimal in general and only leads to about 1% change in the average Dice coefficients for these structures. It is worth noting that even for the Siemens data sets the overall Dice coefficient improves with the atlas renormalization procedure, as shown in the last column of Table I.

We note that for some structures like hippocampus and pallidum, the segmentation accuracy is still lower, in general, for the GE data than the Siemens data even when the atlas renormalization procedure is applied. One explanation is that the GE data were acquired using a different imaging sequence (SPGR) compared to the Siemens scans (MPRAGE) and with slightly lower imaging resolution, which appear to have lower intrinsic image contrast as can be observed from Fig. 4. In addition, for structures like pallidum, the local linear registration is hard to compute accurately due to its small size and reduced contrast with respect to surrounding structures, hence the computed intensity mapping may be less reliable.

V. CONCLUSION

We have developed an atlas renormalization procedure to account for changes in image intensity contrast for an atlas-based automatic brain segmentation method. Testing results on 27 subjects have shown that the new procedure has a minimal effect on the performance of the atlas-based segmentation method when no contrast change is expected but dramatically improves the segmentation accuracy when the same atlas is applied to image data acquired at a different scanner platform (average Dice coefficients improved by 10% or more for several structures). The procedure thus yields more robust segmentations, which can facilitate multicenter neuroscience and clinical studies that typically involve data acquired across different scanner platforms, as well as rendering the tools largely insensitive to scanner upgrades and variations in pulse parameters. In future work, more thorough evaluation will be conducted by acquiring additional test data from other scanner platforms or pulse sequences. Future work also includes further improvement of the accuracy and robustness of the multilinear structure registration, by taking into account spatial correlation of the multiple brain structures and by incorporating additional prior model about plausible local deformations, which can be estimated during atlas construction.

REFERENCES

- [1] X. Zeng, L. H. Staib, R. T. Schultz, and J. S. Duncan, "Segmentation and measurement of the cortex from 3D MR images using coupled surfaces propagation," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 100–111, Oct. 1999.
- [2] X. Han, D. L. Pham, D. Tosun, M. E. Rettmann, C. Xu, and J. L. Prince, "CRUISE: Cortical reconstruction using implicit surface evolution," *NeuroImage*, vol. 23, pp. 997–1012, 2005.
- [3] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis. I. Segmentation and surface reconstruction," *NeuroImage*, vol. 9, pp. 179–194, 1999.
- [4] D. MacDonald, N. Kabani, D. Avis, and A. C. Evans, "Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI," *NeuroImage*, vol. 12, pp. 340–356, 1999.
- [5] D. C. van Essen, J. Dickson, J. Harwell, D. Hanlon, C. H. Anderson, and H. A. Drury, "An integrated software system for surface-based analyses of cerebral cortex," *J. Amer. Med. Info. Assoc.*, vol. 8, pp. 443–459, 2001.
- [6] D. W. Shattuck and R. M. Leahy, "BrainSuite: An automated cortical surface identification tool," *Med. Image Anal.*, vol. 6, pp. 129–142, 2002.
- [7] C. Xu, D. L. Pham, M. E. Rettmann, D. N. Yu, and J. L. Prince, "Reconstruction of the human cerebral cortex from magnetic resonance images," *IEEE Trans. Med. Imaging*, vol. 18, no. 6, pp. 467–480, Jun. 1999.

- [8] D. V. Iosifescu, M. E. Shenton, S. K. Warfield, R. Kikinis, J. Dengler, F. A. Jolesz, and R. W. McCarley, "An automated registration algorithm for measuring MRI subcortical brain structures," *NeuroImage*, vol. 6, pp. 13–25, 1997.
- [9] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 828–839, Oct. 1999.
- [10] P. M. Thompson and A. W. Toga, "Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations," *Med. Image Anal.*, vol. 1, pp. 271–294, 1997.
- [11] B. M. Dawant, S. L. Hartmann, J.-P. Thirion, F. Maes, D. Vandermeulen, and P. Demaerel, "Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations," *IEEE Trans. Med. Imag.*, vol. 10, no. 10, pp. 909–916, Oct. 1999.
- [12] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. M. Dale, "Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain," *Neuron*, vol. 33, pp. 341–355, 2002.
- [13] J. Zhou and J. C. Rajapakse, "Segmentation of subcortical brain structures using fuzzy templates," *NeuroImage*, vol. 28, pp. 915–924, 2005.
- [14] B. Fischl, D. Salat, A. Van der Kouwe, N. Makris, F. Segonne, B. T. Quinn, and A. Dale, "Sequence-independent segmentation of magnetic resonance images," *NeuroImage*, vol. 23, pp. 69–84, 2004.
- [15] J. R. Mitchell, C. Jones, J. Karluk, K. Kennedy, D. H. Lee, B. Rutt, and A. Fenster, "MR multispectral analysis of multiple sclerosis lesions," *J. Magn. Reson. Imag.*, vol. 7, pp. 499–511, 1997.
- [16] L. G. Nyul and J. K. Udupa, "On standardizing the MR image intensity scale," *Magn. Reson. Med.*, vol. 42, pp. 1072–1081, 1999.
- [17] L. G. Nyul, J. K. Udupa, and X. Zhang, "New variants of a method of MRI scale standardization," *IEEE Trans. Med. Imag.*, vol. 19, no. 2, pp. 143–150, Feb. 2000.
- [18] Y. Ge, J. K. Udupa, L. G. Nyul, L. Wei, and R. I. Grossman, "Numerical tissue characterization in MS via standardization of the MR image intensity scale," *J. Magn. Reson. Imag.*, vol. 12, pp. 715–721, 2000.
- [19] A. Madabhushi and J. K. Udupa, "New methods of MR image intensity standardization via generalized scale," *Med. Phys.*, vol. 33, pp. 3426–3434, 2006.
- [20] W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz, "Adaptive segmentation of MRI data," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 429–442, Aug. 1996.
- [21] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "A unifying framework for partial volume segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 22, no. 1, pp. 105–119, Jan. 2003.
- [22] V. Arsigny, X. Pennec, and N. Ayache, "Polyrigid and polyaffine transformations: A novel geometric tool to deal with non-rigid deformations—Application to the registration of histological slices," *Med. Image Anal.*, vol. 9, pp. 507–523, 2005.
- [23] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 1994.
- [24] S. Periaswamy and H. Farid, "Elastic registration in the presence of intensity variations," *IEEE Trans. Med. Imag.*, vol. 22, no. 7, pp. 865–874, Jul. 2003.
- [25] C. J. van Rijsbergen, *Information Retrieval*, 2nd ed. London, U.K.: Butterworths, 1979.