HIPPOCAMPUS 19:565–571 (2009)

Appearance and Incomplete Label Matching for Diffeomorphic Template Based Hippocampus Segmentation

John Pluta,[1,2]* Brian B. Avants,[1,2] Simon Glynn,[1,2] Suyash Awate,[1,2] James C. Gee,[1,2] and John A. Detre[1,2]

ABSTRACT: We present a robust, high-throughput, semiautomated template-based protocol for segmenting the hippocampus in temporal lobe epilepsy. The semiautomated component of this approach, which minimizes user effort while maximizing the benefit of human input to the algorithm, relies on "incomplete labeling." Incomplete labeling requires the user to quickly and approximately segment a few key regions of the hippocampus through a user-interface. Subsequently, this partial labeling of the hippocampus is combined with image similarity terms to guide volumetric diffeomorphic normalization between an individual brain and an unbiased disease-specific template, with fully labeled hippocampi. We solve this many-to-few and few-to-many matching problem, and gain robustness to inter and intrarater variability and small errors in user labeling, by embedding the template-based normalization within a probabilistic framework that examines both label geometry and appearance data at each label. We evaluate the reliability of this framework with respect to manual labeling and show that it increases minimum performance levels relative to fully automated approaches and provides high inter-rater reliability. Thus, this approach does not require expert neuroanatomical training and is viable for high-throughput studies of both the normal and the highly atrophic hippocampus. © **2009 Wiley-Liss, Inc.**

KEY WORDS: hippocampus; normalization; segmentation

INTRODUCTION

In mesiotemporal lobe epilepsy, atrophy as a result of hippocampal sclerosis has long been regarded as an excellent biomarker for aiding in the determination of the laterality of seizures (Duzel, 2006). In pharmacologically intractable cases, temporal lobectomy is one of the few treatments remaining that offers a reasonable chance of relief (Duzel, 2006), which makes seizure localization an important determinant in presurgical evaluation. Historically, atrophy has been measured by determining the degree of asymmetry between the two hippocampi, in the form of an asymmetry ratio computed from the measured volumes of each. The principal difficulty in calculating this ratio lies in accurately obtaining those volumes.

There currently exists many ways to obtain hippocampal volumes. Manual segmentation, where a trained operator reviews an anatomical image and denotes voxels belonging to the hippocampus slice by slice, is typically considered the gold standard for hippocampal measurement (Chupin, 2007). The principle drawback to manual segmentation is the extensive time required of the operator, both in training and time spent performing the actual segmentations- usually a minimum of 30–40 min (and often over an hour) per hippocampus for an experienced user. This can quickly escalate to a nontrivial time commitment in even a small data set. It is also highly dependent on the consistent training among raters, and is most subject to both inter and intrarater bias (Chupin, 2007).

Several automatic methods exist to perform segmentation more quickly. Commonly, a normalization is performed between each subject and a predefined template, and then a hippocampus defined in template space is warped back into the native space of the subject. Other methods instead use shape matching and boundary definitions to perform the whole segmentation in native space. While these methods have been shown to create accurate segmentations of normal, healthy subjects, they are not suited to the segmentation of anatomy that has been affected by disease, injury, aging, or some other condition.

Semiautomated algorithms attempt to make segmentation of abnormal structures possible by having some limited involvement from the user, typically in the form of placing landmarks to guide the creation of the hippocampal mask. However, many of the existing methods that have achieved a reasonable degree of agreement with manual segmentations often require an impractical amount of landmarks, up to 200 per hippocampus (Shen, 2002). This demands not only a time investment, but significant training in hippocampal anatomy. There are methods that employ fewer landmarks (Hogan, 2000), but have consequently less accurate results.

In this article, we present and evaluate a freely available tool that grants the user advantages from this range of methods (Fig. 1). Our approach provides a custom, multisubject template, and requires only that the user do a partial manual segmentation by placing approximate labels in regions of the hippocampus that are likely to pose difficulty for normalization. This user-based partial labeling (described in section

[1] **Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania;** [2] **Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania**
*Correspondence to: John Pluta, Departments of Radiology and Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA.
E-mail: jpluta@mail.med.upenn.edu
Accepted for publication 11 March 2009
DOI 10.1002/hipo.20619
Published online 12 May 2009 in Wiley InterScience (www.interscience.wiley.com).

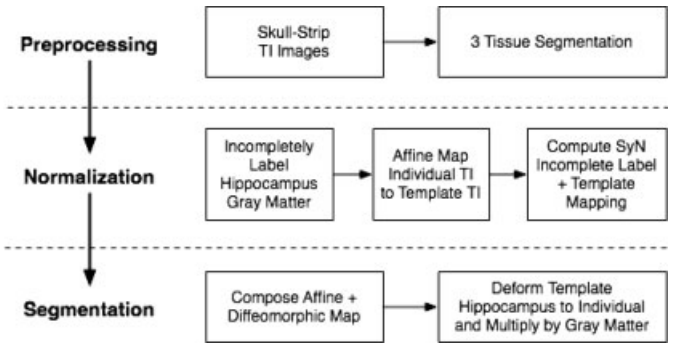

FIGURE 1. The incomplete labeling protocol follows a series of steps to guarantee ease and consistency of labeling and reliable performance even when the user lacks expert neuroanatomical knowledge.

Incomplete Labeling Protocol) is compared to an existing complete hippocampus labeling that exists in the template space through a diffeomorphic normalization based on the labels and image similarity over the whole brain. A probabilistic estimate of the matching between the complete and partial set of labels relieves the many-to-few, under-constrained nature of the problem. Solving the partial labeling correspondence problem, rather than the explicit landmarking problem, reduces the effect of inter and intra rater variance in the placement of the labels, thus greatly increasing the stability of segmentation estimates and derived quantities. Furthermore, we demonstrate that this method is effective and consistent both on normal and atrophic hippocampi.

METHODS

Preprocessing

A template was generated out of the T1-anatomical images of all 19 patients using unbiased shape and appearance averaging in the diffeomorphic space (Kim, 2008). Our template construction process is similar to that used in (Joshi, 2004), but contains an explicit geometric component that directly minimizes the shape metric. Disease was confined to one side by flipping all of the subjects with right-sided epilepsy and then combining the whole group. A complete label of both template hippocampi was created by one of the authors (J.P.).

Incomplete Label Normalization with Diffeomorphisms

We use the transformation group Diff_0, the diffeomorphisms of domain $\Omega \rightarrow \Omega$ with fixed boundary, to represent our anatomical mappings. Define $\phi(x, t) \in \mathrm{Diff}_0$ as a path of diffeomorphisms, where t is a time parameter that controls progress along the path. Without loss of generality, ϕ may be defined as consisting of two components ϕ_1, ϕ_2 such that $\phi(x, 1) = \phi_2^{-1}(\phi_1(x, t), 1 - t)$. This decomposition of ϕ enables symmetric diffeomorphic normalization (SyN) [11], which makes the solution to the correspondence problem unbiased with respect to the input data. In general, an optimal SyN mapping between two images, I and J, computes ϕ_1, ϕ_2 such that $\|I(\phi_1(x, 0.5)) - J(\phi_2(x, 0.5))\|$ is minimized. A change of variables shows that both $\|I(\phi_1(x, 1)) - J(x)\|$ and $\|I(x) - J(\phi^{-1}(x, 1))\|$ are also minimized. Regularity of these mappings are guaranteed by restricting ϕ to be within the diffeomorphic space (the space of differentiable maps with differentiable inverse). The problem is made well-posed by selecting the diffeomorphism with shortest length that maximizes similarity. The length of the diffeomorphism is measured as the integrated length of the velocity fields $v(x, t)$, that generate ϕ through the ordinary differential equation $\frac{d\phi(x,t)}{dt} = v(\phi(x, t), t)$. Note that $v(x, t)$ may in fact be constant in time.

Next, we state the diffeomorphic image matching problem with use of completely and incompletely labeled neuroanatomy. Our approach is novel in that we derive an expected matching term from a probabilistic comparison of a complete segmentation to an incomplete (or partial) segmentation. We also include an expected matching term that derives from probabilistically comparing the incomplete segmentation to the complete segmentation. This bi-directional view of the optimization aids convergence to an accurate hippocampus segmentation, given our specific application. This statement may be shown empirically, but is left to future work. Perhaps the most closely related work, within the field of medical image normalization, is contained in (Guo, 2005), where probability densities are compared directly with a uni-directional optimization. In contrast, we derive expected values from probabilities and exploit the advantages of a bi-directional optimization.

First, define $\{Q\}$ as the set of points describing the complete hippocampus labeling within the template space, I. That is, $Q_i \in \{Q\}$ is the Euclidean position of some labeled point in patient space. Define $\{q\}$ as the set of points describing the incomplete hippocampus labeling within the individual space, J. Here, J is an individual after affine normalization to the template, I. Note that the cardinality of $\{Q\}$, denoted n, and the cardinality of $\{q\}$, m, may not be the same. In most cases, $m \ll n$, as $\{q\}$ is the incomplete label set. After normalization, $\{q\}$ will map to a subset of $\{Q\}$.

One may define the probability of a point-set $\{Q\}$ with respect to a point in $\{q\}$ using a nonparametric Parzen windowing approach. Note that $\Pr(\{Q\}|q_j) = \Pr(q_j|\{Q\})\Pr(\{Q\})$. The prior term, $\Pr(\{Q\})$, is a constant in this application, but may prove useful in future work. The probabilistic relationship between $\{Q\}$ and q_j, in terms of some feature function, f: $(Q_i, q_j) \rightarrow \mathbb{R}^+$ is then,

$$\Pr(\{Q\}|q_j) = \sum_{i=1}^{i=n} W \; Exp\left(-\frac{f(Q_i, q_j)}{\sigma_Q}\right), \qquad (1)$$

where W is a scalar that normalizes the density to sum to one and σ_Q is a scalar variance estimate. We compute the same type of relationship between some Q_i and the set $\{q\}$ with variance parameter σ_q and the same feature function. Here, $\sigma_Q =$

$\sigma_q = 4$ mm; this value was chosen empirically. For this analysis, we choose f to be the squared Euclidean distance between the two points q_j and Q_i. Below, $\Pr(\phi_1(\{Q\})|\phi_2(q_j)) = \Pr(\phi_1(\{Q\}))_{\phi_2 q_j}$ denotes the probability of a point-set $\{Q\}$ mapped by ϕ_1 at 0.5 taken with respect to the set q_j mapped by ϕ_2 at 0.5. Similarly for $\Pr(\phi_2(\{q\})_{\phi_1,Q_i}$.

Denote the expectation of a point-set, given some prior data, as $E(\{Q\})_{q_j} = \sum_i \Pr(Q_i|q_j)Q_i = \sum_i W\ Exp\left(-\frac{1f(Q_i,q_j)}{\sigma_Q}\right)Q_i$. The method below minimizes the distance between each point-set and the expected corresponding point in the opposing point-set.

Diffeomorphic **M**atching **W**ith **I**ncomplete **L**abeling:

Given *I, J,* {*q*}, {*Q*}, and some initial estimate for ϕ_1, ϕ_2, find the shortest $\phi \in$ Dif f_0 such that:

1. $||I(\phi_1(x,0.5)) - J(\phi_2(x,0.5))||$ is minimal
2. $\sum_j w_q^j||\phi_2(q_j,0.5) - E(\phi_1(\{Q\},0.5))_{\phi_2 q_j}||$ is minimal and
3. $\sum_i w_Q^i||\phi_1(Q_i,0.5) - E(\phi_2(\{q\},0.5))_{\phi_1 Q_i}||$ reaches a constrained minimum

This translates to the following variational problem, with linear operator L,

$$
\begin{aligned}
&E_{\mathrm{SyN}}(I,J,\{q\},\{Q\}) \\
&\quad = \inf_{\phi_1}\inf_{\phi_2}\int_{t=0}^{0.5}\{\langle Lv_1,v_1\rangle + \langle Lv_2,v_2\rangle\}dt \\
&+ \int_\Omega [w|I(\phi_1(x,0.5)) - J(\phi_2(x,0.5))|^2 + \sum_j w_q^j||\phi_2(q_j,0.5) \\
&\quad - E(\phi_1(\{Q\},0.5))_{\phi_2 q_j}||^2 + \sum_i w_Q^i||\phi_1(Q_i,0.5) \\
&\quad - E(\phi_2(\{q\},0,5))_{\phi_1 Q_i}||^2]d\Omega
\end{aligned}
$$

with w a scalar and w_q and w_Q scalar, spatially varying weights and each ϕ_i the solution of:

$$\frac{d\phi_i(x,t)}{dt} = v_i(\phi_i(x,t),t) \text{ with } \phi_i(x,0) = x \qquad (2)$$

In this work, we set $w = 1$ as a constant, w_q^j is a Gaussian smoothed dirac delta function centered at each deformed, labeled point with value 3 (chosen to weight the label force strongly relative to image forces) at the peak and w_Q^i starts with the same value as w_q^j. We optimize this problem by gradient descent on the total energy, in a multiresolution framework. However, to achieve a constrained optimization of the third term, we relax (that is, reduce) the weighting on the third term over optimization time. This relaxation is performed at each of three resolutions in a dyadic Gaussian image pyramid, with the additional setting that w_Q is very small towards the latter half of the optimization at each resolution. The effect prevents an over-fitting of the complete labeling to the incomplete labeling. Note that the formulation is the same for one hippocampus or two or a variety of other labels.

Additionally, a standard symmetric diffeomorphic normalization (one not guided by partial labels) was performed on each subject to generate a complete hippocampal labeling with no input from the user. This was done to contrast how segmentations generated by a standard normalization with no user input (fully-automatic segmentations) differ from those generated by guided normalization (semiautomatic segmentations).

Manual Segmentation

Manual segmentation and partial labeling were both performed using ITK-SNAP software package (Yushkevich, 2006), which allows concurrent viewing of sagittal, coronal, and axial planes, as well as three-dimensional representation of the current segmentation. These were completed by one of the authors (J.P.) with prior training in hippocampal anatomy and segmentation, using commonly referenced, published anatomical guidelines (Hasboun, 1996) along with neuroanatomical atlases (Duvernoy, 2005) for further boundary definition.

Incomplete Labeling Protocol

Exactly six labels were placed by one of the authors (JP), in both healthy and atrophic hippocampal. These six regions were chosen by compiling manual segmentation data of 10 control hippocampi into a single normal space, and finding the areas with the most variance between segmentations. Then in each subject, the region approximately corresponding to each of these areas of high variance is masked. The process of identifying and labeling these areas was adapted into a protocol that was subsequently applied to each subject. By manually labeling these areas, the user maximizes the efficiency of the segmentation algorithm while minimizing their involvement.

The protocol for incomplete labeling (illustrated in Fig. 2) is defined as follows:

1. The first label is placed at the most lateral point of the hippocampus where the superior boundary of the alveus was first visible, and was constrained by the alveus and parahippocampal gyrus. All of the labels should be mostly within the gray matter of the hippocampus.
2. The second and third labels are placed three slices from the center of the first label, at the head and tail end of the hippocampus.
3. The fourth label is placed at the medial (towards mid-sagittal) boundary of the mid-body of the hippocampus.
4. The fifth and sixth labels are placed at the posterior and anterior boundaries of the hippocampus in the last slice that they are still clearly definable from the surrounding tissue. This is usually within three slices of the fourth label.

Gray matter probability maps for each subject were generated using the FAST (Zhang, 2001) segmentation procedure. Once normalization was complete, the template hippocampus was warped back to each subject's space, multiplied by the gray matter probability map, and thresholded to generate the final segmentation.

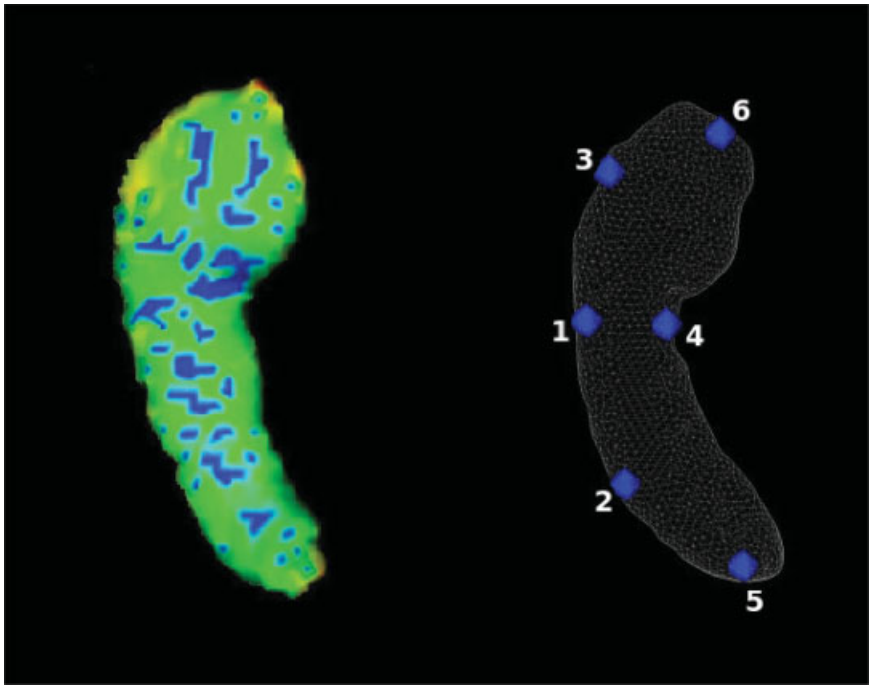

FIGURE 2. Composite map of the variability between hippocampal segmentations (left), and wireframe representation of a hippocampus with labels placed (right). Labels help define the shape of the hippocampus by being manually placed in the areas with the greatest variance. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Data Acquisition

All magnetic resonance imaging (MRI) data presented was acquired from a functional neuroimaging study of refractory temporal lobe epilepsy (TLE). Patients had refractory nonlesional or lesional TLE, while patients with brain tumors, vascular lesions of the temporal lobe, or extratemporal epilepsy were excluded. Furthermore, patients with prior temporal lobectomy and those who were unable to be in an MRI environment were also excluded. High-resolution, T1-weighted anatomical data (voxel sizes of 0.9766 × 0.9766 × 1 mm) was collected for all patients on a 3T Siemens Trio MRI scanner (Siemens, Germany) using a Bruker coil.

The dataset consisted of 19 patients, seven with TLE localized to the right temporal lobe, 10 with TLE localized to the left temporal lobe, and two with bilateral TLE. The patients ranged in age from 18 to 67, with a mean age of 40.0 (SD = 13.5). Sixteen of the participants were right-handed, three were left handed (one localized to the right side and two localized to the left side), and one was ambidextrous and localized to the left side. Sixteen participants were women and four were men. One subject was removed because of an imaging artifact that made automated segmentation impossible.

RESULTS

Volumes of the hippocampal segmentations produced by each method, for both healthy and atrophic subjects, are displayed in Figure 3. Automatic segmentation displays a strong tendency to overestimate the volume of the hippocampus, and shows much more variability than label guided segmentation. While the correlation between manual segmentation and both methods is high, partial labeling provides a superior volume correlation, even in atrophic cases. Standard normalization was able to capture the lower volumes present in atrophy, but had a tendency to overestimate volumes in both cases.

The validities of the fully automated and semiautomated methods were further quantified by calculating the overlap percentage against the gold-standard manual segmentation for each subject. This was done using a Dice Similarity Coefficient; given two segmentations, X and Y, the coefficient is calculated as

$$D = \frac{2|X \cap Y|}{|X| + |Y|} \tag{3}$$

Figure 4 shows the Dice overlap between each subject's manual segmentation and the segmentations generated by both the fully-automated and label-based methods. In healthy hippocampi, the results are quite similar, with a small difference in variance. The improvement is more pronounced in the sclerotic hippocampi, where labeling offers a substantial increase in accuracy and consistency. Most important is the lack of outliers, implying that the guaranteed minimum performance is high.

Table 1 presents the volume, in voxels, of atrophic hippocampi from each subject, as generated by each method of segmentation (mean overlap for controls was 0.8364 for M vs. A

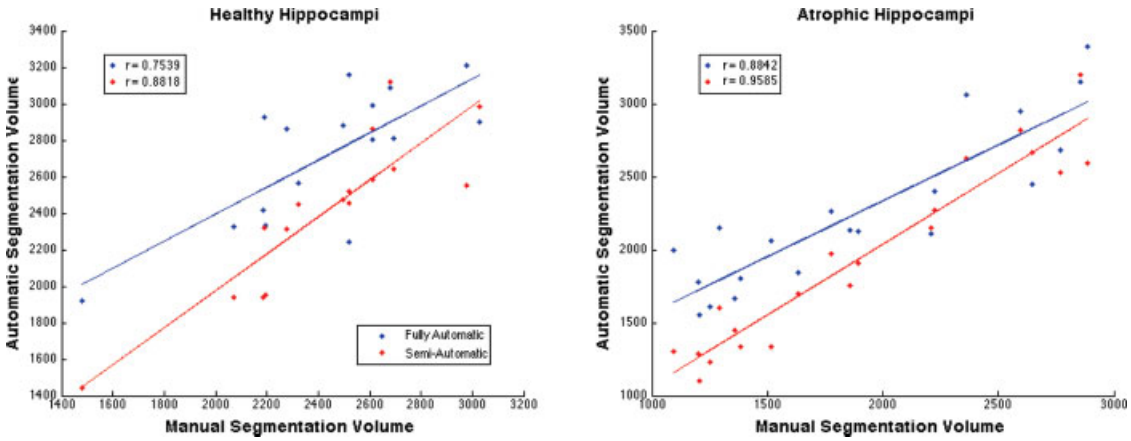

FIGURE 3. Graphs of hippocampal volumes produced by manual segmentation vs. volumes produced by fully and semiautomated segmentation. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

and 0.8649 for M vs. SA). Then, a Dice coefficient was calculated between the manual segmentation and the fully-automatic and semiautomatic methods. Both methods demonstrate an exceptional degree of agreement with the human segmentation. Initialization with labels produced comparable results as with automatic segmentation, but the main benefit is greatly reduced variability, as is demonstrated by the minimal variance. Note that for subjects 10 and 14, who had bilateral TLE, both hippocampi were included in the diseased group.

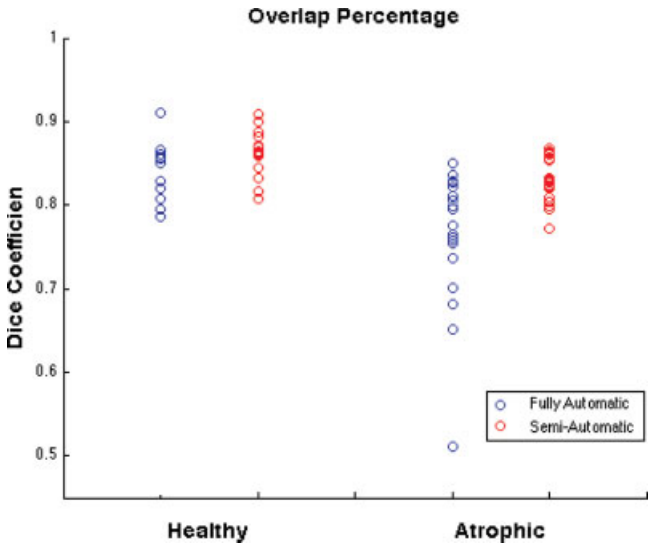

FIGURE 4. Graphs of the Dice coefficients between manual and fully-automated, and manual and semiautomated segmentations for healthy and atrophic hippocampi. Results are comparable for controls, with partial labeling incurring less variance and producing a slightly higher mean. For patients, partial labeling substantially increases performance while decreasing variability. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Finally, we contrast the overlap between segmentations generated by incomplete label-based normalization as determined from labels set by two different users. Due to space considerations, we report only the mean and standard deviation (SD)

TABLE 1.

Volume (in Voxels) and Overlap Percentages of Diseased Hippocampi as Generated by Expert Manual Segmentation, Automatic Segmentation (no Labels Placed), and Semiautomatic Segmentation (Labels Placed)

	Atrophic hippocampi				
Subject	Manual	Automatic	Semiautomatic	M vs. A	M vs. SA
S02 R	2,769	2,678	2,532	0.762254	0.822486
S03 L	2,226	2,403	2,273	0.850724	0.853968
S04 R	2,648	2,446	2,668	0.822144	0.855154
S05 R	1,253	1,612	1,230	0.753927	0.819976
S06 L	1,777	2,263	1,974	0.775743	0.809384
S07 L	2,359	3,064	2,625	0.758621	0.864767
S08 R	1,095	1,993	1,301	0.680699	0.820534
S09 R	1,383	1,806	1,331	0.795861	0.867354
S10 L	1,202	1,778	1,290	0.765101	0.833066
S10 R	1,519	2,061	1,337	0.811732	0.859944
S11 L	2,209	2,110	2,147	0.805279	0.827824
S13 L	1,207	1,555	1,099	0.737147	0.795317
S14 L	2,597	2,946	2,820	0.826989	0.860255
S14 R	2,885	3,389	2,590	0.836468	0.862466
S15 R	1,291	2,147	1,603	0.651593	0.771207
S16 L	1,634	1,839	1,698	0.761301	0.822929
S17 L	1,894	2,127	1,911	0.799304	0.803679
S18 R	1,858	2,134	1,752	0.700902	0.829917
S19 L	1,358	1,662	1,450	0.511258	0.799145
S20 L	2,769	2,678	2,532	0.762254	0.822486
		Mean		0.76182595	0.83041225
		Standard Deviation		0.07913747	0.02686735

Dice overlap for consistency between the two raters, over 10 hippocampi, as 0.895567 ± 0.03%. This high degree of agreement implies that the segmentations generated are very similar, despite partial labels being placed by two different raters and presumably in slightly different locations. Thus, this technique reduces the impact of user decisions, and the incomplete labeling approach greatly improves standard inter-rater overlap while reducing the human time involved for producing full hippocampus segmentations by a factor of twenty or more.

DISCUSSION

In this article, we have presented early work on a robust method for accurate semiautomated segmentation of both normal and abnormal hippocampi. We have demonstrated the capability to produce specific hippocampal masks that approach the same quality as a segmentation done by hand, with a minimal amount of variance, in cases with varying degrees of sclerosis. All of the software used in this article is freely available at http://www.picsl.upenn.edu/ANTS, and includes a tutorial detailing usage, a guide to label placement, and a primer on hippocampal anatomy.

Chupin et al. (2007) provides a comprehensive evaluation of the performance of current semiautomated methods for abnormal hippocampi, complete with mean overlap measures, processing time, and manual interaction time. Mean overlap, inter-rater reliability, and intrarater reliability, and volume measurements are comparable between the method we present and the current literature. We note that firm conclusions cannot be made by comparing performance values derived from different datasets and with likely different definitions of what is and is not hippocampus.

Cases with the lowest performance, such as s15, were typically subjects with the most severe degree of atrophy. In these instances, the subject's anatomy deviated greatly even from the disease-specific template, and made normalization more difficult. Mismatching of the template hippocampus to some other nearby structure was a problem in these cases, though this was due the subjects anatomy no longer resembling the template. This was much more of a problem in automatic segmentation, where there were no labels to guide normalization. Performance was consequently lower.

Volume measurements showed a high correlation with manual segmentation volumes ($r = 0.8818$ for healthy hippocampi, $r = 0.9585$ for atrophic hippocampi), and this correlation is largely maintained even in subjects where the Dice coefficient was lower than average. Semiautomatic segmentation predicted the same asymetries as manual segmentation in every case except for s4, s7, and s14. However, in all of these cases the volumes of both hippocampi were quite close to each other and there was not sufficient atrophy evident to be used in determining seizure laterality. Extreme outliers pose a significant problem to this and any normalization method Hammers (2007) and represents a major limitation that is only somewhat mitigated by partial labeling.

The principle advantage of our method is that it requires only a minimal time investment from the user, both in training and manual delineation. A detailed knowledge of neuroanatomy is not required and only six labels are used in the labeling protocol, thus a practiced user can label a dataset (both hippocampi of a single subject) in 5 min or less. At the same time, the method also guarantees high values for both expected performance and minimum performance in terms of segmentation quality. Furthermore, the sensitivity of the algorithm to the placement of these labels is minimized by using an incomplete labeling.

This technique has been used to generate hippocampal segmentations in several other abnormal datasets, including Alzheimer's disease, brain injury, and aged controls. Each individual segmentation was briefly reviewed and all were of comparable quality to those presented in this dataset, though the results were not evaluated as rigorously. None of these cases required any adjustment to the partial labeling protocol, or any additional preprocessing beyond the creation of a template and template segmentation.

The focus of future efforts will be on optimizing the normalization methodology to produce even better segmentations. We will explore joint probability frameworks, explicit outlier rejection and additional or alternative similarity terms driving the final result. Another principle interest will be to continue to apply this technique to the segmentation of hippocampi in other abnormal datasets, and other structures of the brain. We intend to generate a user community for this methodology which will likely lead to currently unforeseeable user-guided improvements in methodology and labeling protocol.

REFERENCES

Chupin M, Mukuna-Bantumbakulu A, Hasboun D, Bardinet E, Baillet S, Kinkingnehun S, Lemieux L, Dubois B, Garnero L. 2007. Anatomically constrained region deformation for the automated segmentation of the hippocampus and amygdala: Method and validation on controls and patients with Alzheimer's disease. Neuroimage 34:996–1019.

Duvernoy H. 2005. The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI. Berlin, Germany: Springer.

Duzel E, Schiltz K, Solbach T, Peschel T, Baldeweg T, Kaufmann J, Szentkuti A, Heinze H. 2006. Hippocampal atrophy in temporal lobe epilepsy is correlated with limbic systems atrophy. J Neurol 253:294–300.

Guo H, Rangarajan A, Joshi S. 2005. 3-d diffeomorphic shape registration on hippocampal data sets. Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv 8:984–991.

Hammers A, Heckemann R, Koepp M, Duncan J, Hajnal J, Rueckert D, Aljabar P. 2007. Automatic detection and quantification of hippocampal atrophy on MRI in temporal lobe epilepsy: A proof-of-principle study. Neuroimage 36:38–47.

Hasboun D, Chantôme M, Zouaoui A, Sahel M, Deladoeuille M, Sourour N, Duymes M, Baulac M, Marsault C, Dormont D. 1996. MR determination of hippocampal volume: Comparison of three methods. Am J Neuroradiol 17:1091–1098.

Hogan R, Mark K, Wang L, Joshi S, Miller M, Bucholz R. 2000. Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients. Radiology 216:291–297.

Joshi S, Davis B, Jomier M, Gerig G. 2004. Unbiased diffeomorphic atlas construction for computational anatomy. Neuroimage 23, (Suppl 1):S151–S160.

Kim J, Avants B, Patel S, Whyte J, Coslett B, Pluta J, Detre J, Gee J. 2008. Structural consequences of diffuse traumatic brain injury: A large deformation tensor-based morphometry study. Neuroimage 39:1014–1026.

Shen D, Moffat S, Resnick S, Davatzikos C. 2002. Measuring size and shape of the hippocampus in MR images using a deformable shape model. Neuroimage 15:422–434.

Yushkevich P, Piven J, Hazlett H, GimpelSmith R, Ho S, Gee J, Gerig G. 2006. User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage 31:1116–1128.

Zhang Y, Brady M, Smith S. 2001. Segmentation of brain MR images through a hidden markov random field model and the expectation maximization algorithm. IEEE Trans Med Imaging 20:45–57.