Model-Based Segmentation of Hippocampal Subfields in Ultra-High Resolution In Vivo MRI

Koen Van Leemput[1,2], **Akram Bakkour**[1,3], **Thomas Benner**[1], **Graham Wiggins**[1], **Lawrence L. Wald**[1,4], **Jean Augustinack**[1], **Bradford C. Dickerson**[1,3], **Polina Golland**[2], and **Bruce Fischl**[1,2,4]

[1] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, MGH, Harvard Medical School, USA

[2] Computer Science and Artificial Intelligence Laboratory, MIT, USA

[3] Department of Neurology, MGH, Harvard Medical School, USA

[4] Harvard-MIT Division of Health Sciences and Technology, MIT, USA

Abstract

Recent developments in MR data acquisition technology are starting to yield images that show anatomical features of the hippocampal formation at an unprecedented level of detail, providing the basis for hippocampal subfield measurement. Because of the role of the hippocampus in human memory and its implication in a variety of disorders and conditions, the ability to reliably and efficiently quantify its subfields through in vivo neuroimaging is of great interest to both basic neuroscience and clinical research. In this paper, we propose a fully-automated method for segmenting the hippocampal subfields in ultra-high resolution MRI data. Using a Bayesian approach, we build a computational model of how images around the hippocampal area are generated, and use this model to obtain automated segmentations. We validate the proposed technique by comparing our segmentation results with corresponding manual delineations in ultra-high resolution MRI scans of five individuals.

1 Introduction

Models of brain structures generated from magnetic resonance imaging (MRI) data have grown in complexity in recent years, evolving from simple models with few classes such as gray matter, white matter and cerebrospinal fluid (CSF) [1,2,3], into more complex ones representing a multitude of neuroanatomical structures [4,5,6]. Still, while many brain structures such as the thalamus, the amygdala, or the hippocampus consist of multiple distinct, interacting subregions, they are mostly treated as a single entity because of the limited image resolution of typical structural MRI scans. Recently, however, substantial developments in MR data acquisition technology have made it possible to acquire images with remarkably higher resolution and signal-to-noise ratio than was previously attainable [7]. Such scans show many cortical and subcortical structures in unprecedented detail, and offer new opportunities for explicitly quantifying individual subregions, rather than their agglomerate, directly from in vivo MRI data.

Analyzing large imaging studies of ultra-high resolution MRI scans requires computational techniques to automatically extract information from the images. This is technically difficult because, although the images show greater anatomical detail than traditional MRI scans, many boundaries between substructures of interest remain hard to discern. In manual delineations, the extent of specific subregions is often inferred from the extent of other, more clearly defined structures by relying on prior neuroanatomical knowledge, rather than on local intensity

information alone. The success of automated methods therefore depends critically on computational models that provide prior information about the relative location, shape, and appearance of the structures of interest.

In this paper, we present an automated segmentation technique for the subfields of the hippocampus in ultra-high resolution MRI data based on state-of-the-art computational models. Although the methodology is applicable to other brain structures as well, we identified the hippocampus as our driving application because it is a necessary component in a variety of memory functions, as well as the locus of structural change in aging, Alzheimer's disease (AD), schizophrenia, and other conditions. Distinct hippocampal subregions have been shown to be implicated in different memory subsystems [8,9] and be differentially affected in aging and AD [10]. Therefore, the ability to measure, through in vivo neuroimaging, subtle changes in these subregions promises to have widespread application in both basic neuroscience and clinical research.

2 Model-Based Hippocampal Subfield Segmentation

We use a Bayesian modeling approach, in which we first build an explicit computational model of how an MRI image around the hippocampal area is generated, and subsequently use this model to obtain fully automated segmentations. The model incorporates a *prior* distribution that makes predictions about where neuroanatomical labels typically occur throughout the image, and is based on the generalization of probabilistic atlases [2,3,4,5,11] developed in [12]. The model also includes a *likelihood* distribution that predicts how a label image, where each voxel is assigned a unique neuroanatomical label, translates into an MRI image, where each voxel has an intensity.

2.1 Prior: Mesh-Based Probabilistic Atlas

Let $L = \{l_i, i = 1,...,\}$ be a label image with a total of I voxels, with $l_i \in \{1,...,K\}$ denoting the one of K possible labels assigned to voxel i. Our prior models this image as being generated by the following process:

A (irregular) tetrahedral mesh covering the image domain of interest is defined by the *reference position* of its N mesh nodes $x^r = \{x_n^r, n=1,\ldots,N\}$, and by a set of label probabilities $\alpha = \{\alpha_n, n = 1,...,N\}$. Node n is associated with a probability vector $\alpha_n = \{\alpha_n^1, \ldots, \alpha_n^K\}$, satisfying $\alpha_n^k \geq 0$ and $\sum_k^K \alpha_n^k = 1$, that governs how frequently each label occurs around that node.

The mesh is deformed from its reference position by sampling from a Markov random field (MRF) model regulating the position of the mesh nodes:

$$p(x) \propto \exp\left(-U_{x^r}(x)\right) = \exp\left(-\sum_{t=1}^{T} U_{x^r}^t(x)\right),$$

(1)

where $U_{x^r}^t(x)$ is a penalty for deforming tetrahedron t from its shape in the reference position x^r, and $U_{x^r}(x)$ is an overall deformation penalty obtained by summing the contributions of all T tetrahedra in the mesh. We use the penalty proposed in [13], which goes to infinity if the Jacobian determinant of any tetrahedron's deformation approaches zero, and therefore insures that the mesh topology is preserved.

In the deformed mesh with position x, the probability of observing label k in a pixel i with location x_i is modeled by

$$p_i (k|x) = \sum_{n=1}^{N} \alpha_n^k \phi_n (x_i) , \tag{2}$$

where $\varphi_n(\cdot)$ denotes an interpolation basis function attached to mesh node n that has a unity value at the position of the mesh node, a zero value at the outward faces of the tetrahedra connected to the node and beyond, and a linear variation across the volume of each tetrahedron. Assuming conditional independence of the labels between voxels given the mesh node locations, we obtain the probability of seeing label image

$$L : p(L|\boldsymbol{x}) = \prod_{i=1}^{I} p_i (l_i|\boldsymbol{x}).$$

It has previously been demonstrated that the mesh's connectivity, reference position $\boldsymbol{x}^r$, and label probabilities $\boldsymbol{\alpha}$ can be learned from a set of manually labeled example images [12]. The learning involves selecting the model that maximizes the probability of observing the example label images, or, equivalently, that minimizes the number of bits needed to encode them. An example of the prior, derived from 4 manually labeled hippocampi, is shown in figure 2. Note that the image domain is non-uniformly sampled, with areas containing little information covered by larger tetrahedra.

2.2 Likelihood: Imaging Model

For the likelihood distribution, we employ a simple, often-used model according to which a Gaussian distribution with mean μ_k and variance σ_k^2 is associated with each label k. Given label image L, an intensity image $Y = \{y_i, i = 1,...,I\}$ is generated by drawing the intensity in each voxel independently from the Gaussian distribution associated with its label:

$$p(Y|L, \theta) = \prod_{i=1}^{I} p\left(y_i|\mu_{l_i}, \sigma_{l_i}^2\right) = \prod_{i=1}^{I} \frac{1}{\sqrt{2\pi\sigma_{l_i}^2}} \exp\left(-\frac{(y_i - \mu_{l_i})^2}{2\sigma_{l_i}^2}\right),$$

where the parameters $\theta = \left\{\mu_1, \sigma_1^2, \ldots, \mu_K, \sigma_K^2\right\}$ are assumed to be governed by a uniform prior: $p(\boldsymbol{\theta}) \propto 1$.

2.3 Model Parameter Estimation

In a Bayesian setting, assessing the Maximum A Posteriori (MAP) parameter values $\left\{\widehat{\boldsymbol{x}}, \widehat{\theta}\right\}$ involves maximizing

$$p(x, \theta|Y) \propto p(Y|x, \theta)\, p(x)\, p(\theta) \propto \left(\prod_{i=1}^{I} \sum_{k=1}^{K} p\left(y_i|\mu_k, \sigma_k^2\right) p_i(k|x)\right) p(x) ,$$

which is equivalent to minimizing

$$\sum_{i=1}^{I} \left(-\log\left[\sum_{k=1}^{K} p\left(y_i|\mu_k, \sigma_k^2\right) p_i(k|x)\right]\right) - \log p(x) . \tag{3}$$

We use an EM-style majorization technique [14,15], where we calculate a statistical classification that associates each voxel with each of the neuroanatomical labels

$$W_i^k = \frac{p\left(y_i|\mu_k, \sigma_k^2\right) p_i\left(k|x\right)}{\Sigma_{k'} p\left(y_i|\mu_{k'}, \sigma_{k'}^2\right) p_i\left(k'|x\right)}$$

and subsequently use this classification to construct an upper bound to eq. (3) that touches it at the current parameters estimates [16]:

$$\sum_{i=1}^{I}\left(-\log\left[\prod_{k=1}^{K}\left(\frac{p\left(y_i|\mu_k, \sigma_k^2\right) p_i\left(k|x\right)}{W_i^k}\right)^{W_i^k}\right]\right) - \log p\left(x\right). \tag{4}$$

Optimizing this upper bound w.r.t. the Gaussian distribution parameters $\boldsymbol{\theta}$, while keeping $\boldsymbol{x}$ fixed, yields the closed-form expressions

$$\mu_k = \frac{\sum_{i=1}^{I} W_i^k y_i}{\sum_{i=1}^{I} W_i^k}, \quad \sigma_k^2 = \frac{\sum_{i=1}^{I} W_i^k (y_i - \mu_k)^2}{\sum_{i=1}^{I} W_i^k}, \quad k=1,\ldots,K.$$

With these estimates of $\boldsymbol{\theta}$, the classification and the corresponding upper bound are updated, and the estimation of $\boldsymbol{\theta}$ is repeated, until convergence. We then re-calculate the upper bound, and optimize it w.r.t. the mesh node positions $\boldsymbol{x}$, keeping $\boldsymbol{\theta}$ fixed. Optimizing $\boldsymbol{x}$ is a registration process that deforms the atlas mesh towards the current classification, similar to the schemes proposed in [5,17]. The gradient of eq. (4) with respect to $\boldsymbol{x}$ is given in analytical form through the interpolation model of eq. (2) and the deformation model of eq. (1). We perform this registration by gradient descent. Subsequently, we repeat the optimization of $\boldsymbol{\theta}$ and $\boldsymbol{x}$, each in turn, until convergence.

2.4 Image Segmentation

Once we have an estimate of the model parameters $\left\{\widehat{\boldsymbol{x}}, \widehat{\theta}\right\}$, we can use it to obtain an approximation to the MAP anatomical labeling. Approximating

$$p\left(L|Y\right) = \int_{\boldsymbol{x}}\int_{\theta} p\left(L|Y, \boldsymbol{x}, \theta\right) p\left(\boldsymbol{x}, \theta|Y\right) \mathrm{d}\boldsymbol{x}\mathrm{d}\theta \quad \text{by} \quad p\left(L|Y, \widehat{\boldsymbol{x}}, \widehat{\theta}\right) \propto p\left(Y|L, \widehat{\theta}\right) p\left(L|\widehat{\boldsymbol{x}}\right), \text{we have}$$

$$\widehat{L} = \underset{L}{\arg\max}\, p\left(L|Y\right) \simeq \underset{\{l_i, i=1,\ldots,I\}}{\arg\max} \prod_{i}^{I} p\left(y_i|\widehat{\mu}_{l_i}, \widehat{\sigma}_{l_i}^2\right) p_i\left(l_i|\widehat{x}\right),$$

which is obtained by assigning each voxel to the label with the highest posterior probability, i.e., $\widehat{l}_i = \arg\max_k W_i^k$.

3 Experiments

We performed experiments on ultra-high resolution MRI data collected as part of an ongoing imaging study assessing the effects of normal aging and AD on brain structure. Using a prototype custom-built 32-channel head coil with a 3.0T Siemens Trio MRI system [7], we

acquired images via an optimized high-resolution MPRAGE sequence that enables 380 μm in-plane resolution (TR/TI/TE = 2530/1100/5.39 ms, FOV=448, FA = 7°, 208 slices acquired coronally, thickness = 0.8mm, acquisition time = 7.34 min). To increase the signal-to-noise ratio, 5 acquisitions were collected and motion-corrected to obtain a single resampled (to 380 μm isotropic) high contrast volume that covers the entire medial temporal lobe.

Using a protocol developed specifically for this purpose, the subfields of the right hippocampus were manually delineated in images of 5 subjects (2 younger and 3 older cognitively normal individuals). These delineations included the fimbria, presubiculum, subiculum, CA1, CA2/3, and CA4/DG fields, as well as choroid plexus, hippocampal fissure, and lateral ventricle, as shown in figure 1. Voxels outside of these structures were automatically labeled as gray matter, white matter, or CSF using an EM-based tissue classifier [2].

We restricted our automated analysis to a region of interest (ROI) around the right hippocampus only. To this end, we defined a cuboid ROI of size 100×60×160 voxels in an image of a younger normal individual not included in the study (template image). This ROI was automatically aligned to each image under study using an affine Mutual Information based registration technique [18,19], by first aligning the whole template image covering the entire brain, followed by a registration of the ROI only. Atlas meshes were then computed and applied in the area covered by this ROI in each image.

We used a 3-level multi-resolution optimization strategy, in which the image under study and the atlas mesh were subject to a gradually decreasing amount of spatial smoothing. In order to simplify the optimization process, we restricted the number of labels throughout the multi-resolution scheme to four, merging the gray matter with the presubiculum, subiculum, and CA fields, the white matter with the fimbria, the lateral ventricle with choroid plexus, and CSF with the hippocampal fissure. This restriction was then removed to obtain the final segmentation. The whole segmentation process was fully automated and took about 1.5 hours per subject on a 2.33GHz Intel Core2 processor.

We evaluated our automated segmentation results using a leave-one-out cross-validation strategy: we built an atlas mesh from the delineations in 4 subjects, and used this to segment the image of the remaining subject. We repeated this process for each of the 5 subjects, and compared the automated segmentation results with the corresponding manual delineations. Towards the tail of the hippocampus, the manual delineations no longer discerned between the different subfields, but rather lumped everything together as simply "hippocampus" (see figure 1). Since the starting point of this "catch-all" label was arbitrary chosen in each subject, with its volume ranging from 5 to 17% of the total hippocampal volume in different subjects, voxels that were labeled as such in either the automated or manual segmentation were not included in the comparisons.

For each of seven structures of interest (fimbria, CA1, CA2/3, CA4/DG, presubiculum, subiculum, and hippocampal fissure), we calculated the Dice overlap coefficient, defined as the volume of overlap between the automated and manual segmentation divided by their mean volume. Since we are ultimately interested in detecting *changes* in hippocampal subfields between different patient populations, we also evaluated how well differences in subfield volumes between subjects, as detected by the manual delineations, were reflected in the automated segmentations. To this end, we performed a linear regression on the absolute volumes detected by both methods, calculating Pearson's correlation coefficient for each structure.

4 Results

Figure 1 compares the manual and automated segmentation results qualitatively on a set of cross-sectional slices. The upper half of figure 3 shows the average Dice overlap measure for each of the structures of interest, along with the minimum and maximum across the 5 subjects. All of the larger structures, ranging in average size from 6,100 voxels for CA1 to 14,300 voxels for CA2/3, have an average Dice coefficient of around 0.7 or higher. Smaller structures such as the fimbria (on average 1,700 voxels) and the hippocampal fissure (on average 1,400 voxels) are more challenging and have a lower Dice coefficient of around 0.58 and 0.45, respectively.

The lower half of figure 3 shows, for each structure, the volume differences between the automated and manual segmentations relative to their mean volumes. Regarding Pearson's correlation coefficient, the automatically calculated volumes of CA4/DG and CA2/3 are strongly correlated with the manual ones, with a correlation coefficient of approximately 0.98 ($p \leq 0.004$) and 0.93 ($p \leq 0.024$), respectively. CA1 and subiculum correlate to some degree (correlation coefficient of 0.73 and 0.71, p-values not significant), whereas presubiculum and fimbria do not seem to correlate at all (correlation coefficient of 0.02 and −0.18, p-values not significant). Interestingly, despite the hippocampal fissure's low Dice overlap measure, its automated measurements correlate better with the manual ones than do some structures with much higher Dice coefficients (correlation coefficient 0.85, $p \leq 0.068$). The low Dice coefficient is apparently caused by a systematic underestimation of the hippocampal fissure volume by the automated method.

5 Discussion

In this paper, we demonstrated a model-based approach to automated hippocampal subfield segmentation in ultra-high resolution MRI and presented preliminary results on a small number of subjects. Future work will include a more thorough validation, using more subjects with repeat scans and manual delineations by different raters, so that the accuracy and repeatability of our method can be placed in context. Furthermore, in order to analyze invaluable existing imaging studies that were acquired at more standard image resolutions, we also plan to develop a modified likelihood for standard resolution images that includes an explicit model of the partial volume effect [20].

Acknowledgments

References

1. Wells W, et al. Adaptive segmentation of MRI data. IEEE Transactions on Medical Imaging 1996;15 (4):429–442. [PubMed: 18215925]
2. Van Leemput K, et al. Automated model-based tissue classification of MR images of the brain. IEEE Transactions on Medical Imaging 1999;18(10):897–908. [PubMed: 10628949]
3. Ashburner J, Friston K. Unified segmentation. NeuroImage 2005;26:839–885. [PubMed: 15955494]
4. Fischl B, et al. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33:341–355. [PubMed: 11832223]
5. Pohl K, et al. A Bayesian model for joint segmentation and registration. NeuroImage 2006;31:228–239. [PubMed: 16466677]

6. Heckemann R, et al. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage 2006;33(1):115–126. [PubMed: 16860573]

7. Wiggins G, et al. 32-channel 3 Tesla receive-only phased-array head coil with soccer-ball element geometry. Magn. Res. in Medicine 2006;56(1):216–223.

8. Zeineh M, et al. Dynamics of the Hippocampus During Encoding and Retrieval of Face-Name Pairs. Science 2003;299(5606):577–580. [PubMed: 12543980]

9. Gabrieli J, et al. Separate Neural Bases of Two Fundamental Memory Processes in the Human Medial Temporal Lobe. Science 1997;276(5310):264. [PubMed: 9092477]

10. Small S, et al. Evaluating the function of hippocampal subregions with high-resolution MRI in AD and aging. Microsc. Res. and Tech 2000;51(1):101–108. [PubMed: 11002358]

11. Shattuck D, et al. Construction of a 3D probabilistic atlas of human cortical structures. NeuroImage 2008;39:1064–1080. [PubMed: 18037310]

12. Van Leemput, K. Probabilistic brain atlas encoding using Bayesian inference.. In: Larsen, R.; Nielsen, M.; Sporring, J., editors. MICCAI 2006. LNCS. Vol. 4190. Springer; Heidelberg: 2006. p. 704-711.

13. Ashburner, et al. Image registration using a symmetric prior–in three dimensions. Human Brain Mapping 2000;9(4):212–225. [PubMed: 10770230]

14. Dempster AP, et al. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society 1977;39:1–38.

15. Stoica P, Selén Y. Cyclic minimizers, majorization techniques, and the EM algorithm: A refresher. IEEE Signal Processing Magazine 2004:112–114.

16. Minka T. Expectation-Maximization as lower bound maximization. Technical report, MIT. 1998

17. D'Agostino, E., et al. A unified framework for atlas based brain image segmentation and registration.. In: Pluim, JPW.; Likar, B.; Gerritsen, FA., editors. WBIR 2006. LNCS. Vol. 4057. Springer; Heidelberg: 2006. p. 136-143.

18. Wells W, et al. Multi-modal volume registration by maximization of mutual information. Medical Image Analysis 1996;1(1):35–51. [PubMed: 9873920]

19. Maes F, et al. Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging 1997;16(2):187–198. [PubMed: 9101328]

20. Van Leemput K, et al. A unifying framework for partial volume segmentation of brain MR images. IEEE Transactions on Medical Imaging 2003;22(1):105–119. [PubMed: 12703764]
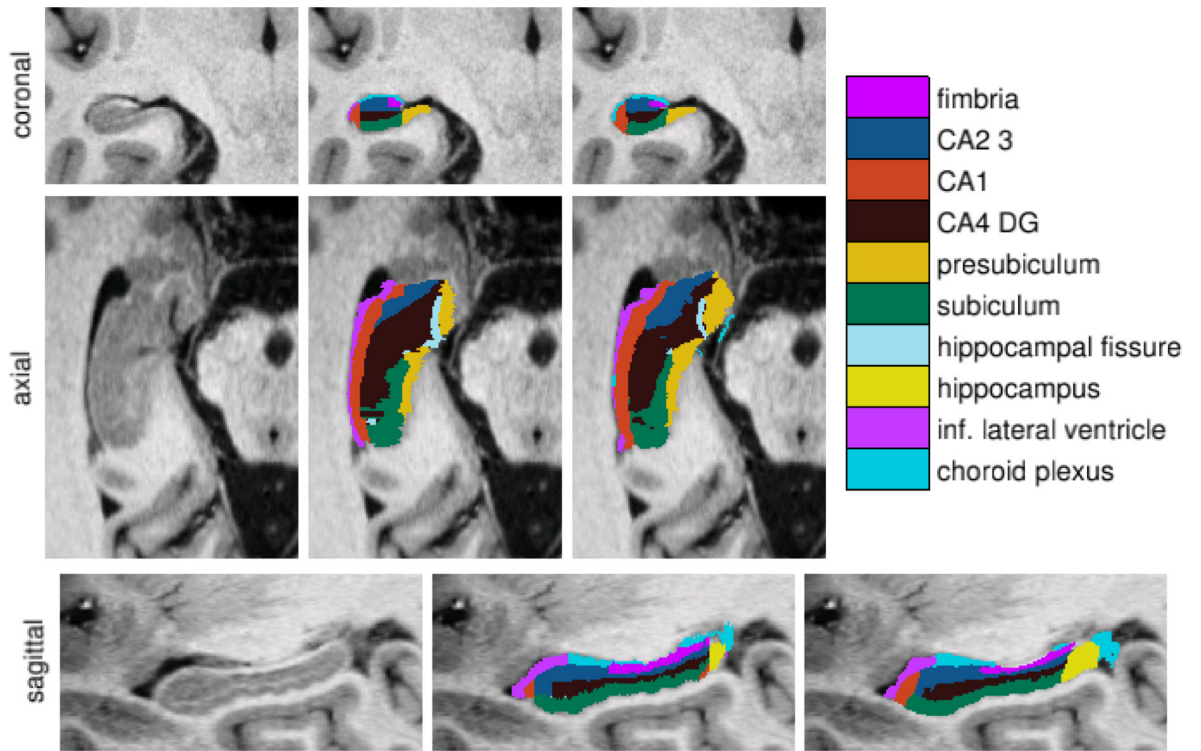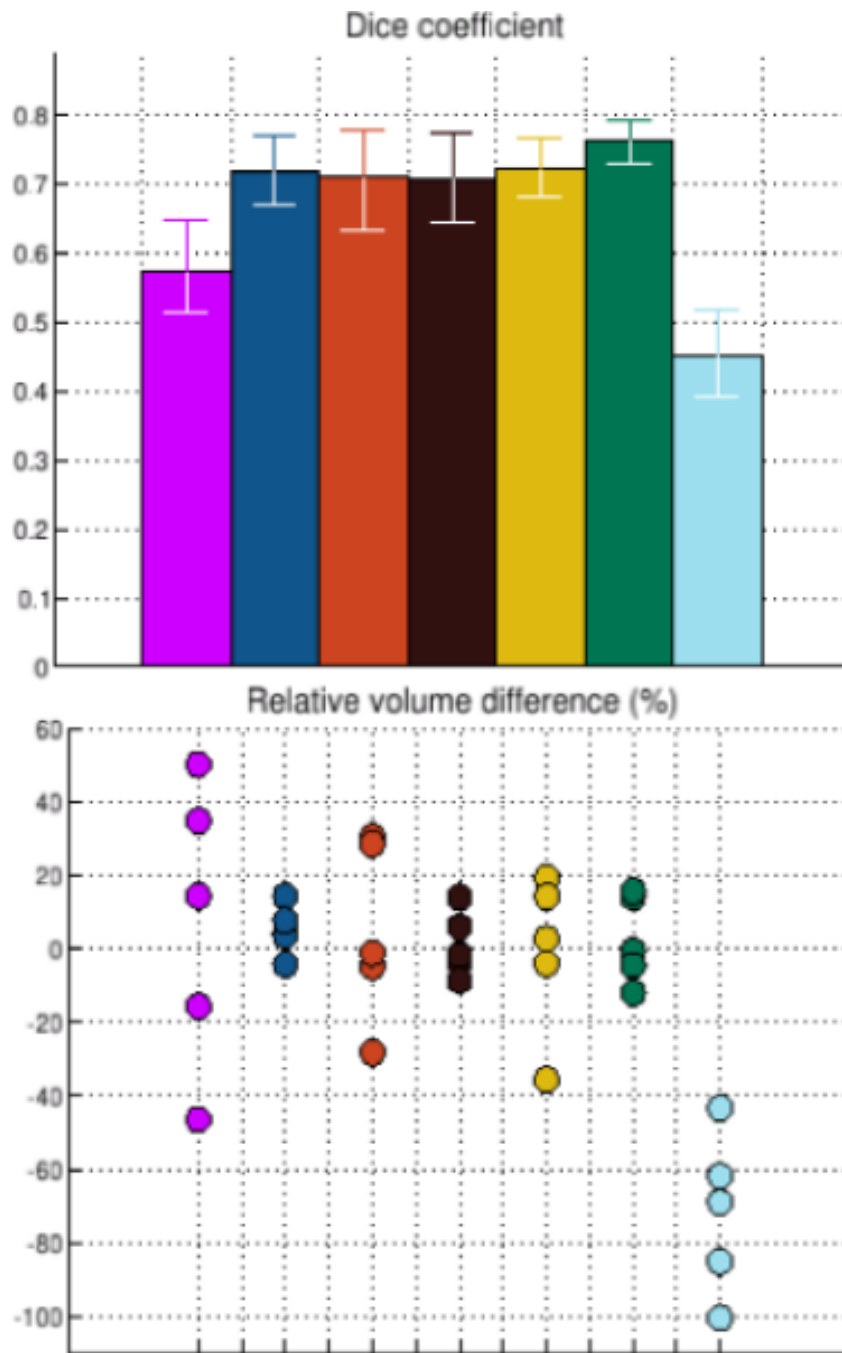
Fig. 1.
From left to right: ultra-high resolution MRI data, manual delineations, and corresponding automated segmentations
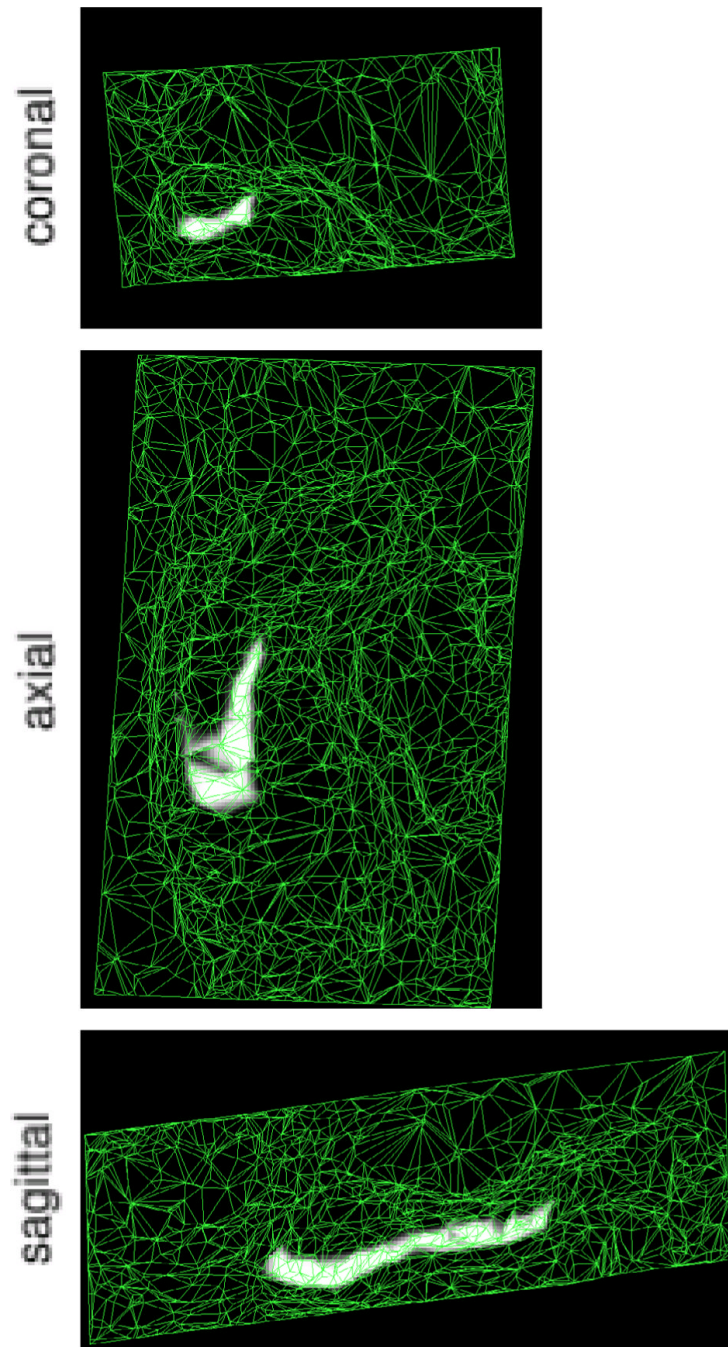
Fig. 2.
Mesh-based probabilistic atlas, derived from manual delineations in 4 subjects, warped onto the 5th subject shown in figure 1. Bright and dark intensities correspond to high and low prior probability for subiculum, respectively.

Fig. 3.
Dice overlap measures (top) and relative volume differences (bottom) between automated and manual segmentations. The colors are as in figure 1.