AUTOMATED 3D MAPPING & SHAPE ANALYSIS OF THE LATERAL VENTRICLES VIA FLUID REGISTRATION OF MULTIPLE SURFACE-BASED ATLASES

Yi-Yu Chou[1], Natasha Leporé[1], Greig I. de Zubicaray[2], Stephen E. Rose[2], Owen T. Carmichael[3], James T. Becker[4], Arthur W. Toga[1], Paul M. Thompson[1]

[1]Laboratory of Neuro Imaging, UCLA Dept. of Neurology, Los Angeles, CA
[2]Univ. of Queensland Center for Magnetic Resonance, Brisbane, Australia
[3]UC Davis, Dept. Neurology, Davis, CA; [4]Univ. Pittsburgh, Dept. Neurology, Pittsburgh, PA

ABSTRACT

We developed and validated a new method to create automated 3D parametric surface models of the lateral ventricles, designed for monitoring degenerative disease effects in clinical neuroscience studies and drug trials. First we used a set of parameterized surfaces to represent the ventricles in a *manually labeled set* of 9 subjects' MRIs (atlases). We fluidly registered each of these atlases and mesh models to a set of MRIs from 12 Alzheimer's disease (AD) patients and 14 matched healthy elderly subjects, and we averaged the resulting meshes for each of these images. Validation experiments on expert segmentations showed that (1) the Hausdorff labeling error rapidly decreased, and (2) the power to detect disease-related alterations monotonically improved as the number of atlases, *N,* was increased from 1 to 9. We then combined the segmentations with a radial mapping approach to localize ventricular shape differences in patients. In surface-based statistical maps, we detected more widespread and intense anatomical deficits as we increased the number of atlases, and we formulated a statistical stopping criterion to determine the optimal value of *N*. Anterior horn anomalies in Alzheimer's patients were only detected with the multi-atlas segmentation, which clearly outperformed the standard single-atlas approach.

1. INTRODUCTION

The ventricular system is a fluid-filled structure in the center of the brain, surrounded by gray and white matter structures that are affected by neurodegenerative diseases. Ventricular measures and surface-based maps provide sensitive assessments of atrophy, which correlate with cognitive deterioration in AD [1,2], HIV/AIDS [3], schizophrenia [4,5] and offer a potential approach to evaluate disease progression in large-scale drug trials.

Despite years of efforts in automated surface parameterization, the concave shape, complex branching topology [6] and extreme narrowness of the inferior and posterior horns have made it difficult for voxel-based classifiers to create accurate ventricular models, and for subsequent surface parameterization approaches to induce a grid on the entire structure [6]. For the vast image datasets now being collected (e.g., 339 subjects in [2]), time and expertise requirements make manual segmentations impractical; more automated methods are urgently needed.

One promising approach nonlinearly registers each image to a target image in which the structures of interest have been labeled and parameterized; the labels are then projected onto the unlabeled data using the computed geometric transformations [7-10]. However, the chosen reference image influences the results of label propagation; different starting atlases result in different segmentations for each subject. In [11-13], multiple segmentations were combined, and the result was closer to manual labels than if only one was used. 'Targetless' image warping also combines multiple pairwise or groupwise registrations, reducing registration errors in groups of images [14].

In this study, we validated a new automated technique for 3D segmentation that combines multiple registration-based surface models into one. We set out to show that this increased label propagation accuracy and the power to detect disease effects, without requiring any interactive human input other than the initial expert labeling of a small set of images. With this method, we then statistically analyzed morphometric differences in the lateral ventricles of subjects with AD, compared to healthy elderly controls. To optimize the algorithm, we examined how the segmentation error (symmetrized Hausdorff distance) and the disease detection power (assessed using cumulative distributions of surface statistics and the False Discovery Rate) depended on the number of atlases.

2. MATERIALS AND METHODS

2.1. Subjects

3D T1-weighted MRI scans of 12 subjects with AD (age: 68.4 $\pm$ 1.9 years) and 14 controls (age: 71.4 $\pm$ 0.9 years) were analyzed. All images were spatially normalized by a 9-parameter transformation (3 translations, 3 rotations, 3

scales) to the International Consortium for Brain Mapping (ICBM-53) average brain template [7].

2.2. Ventricular Delineation

To build the labeled atlases, the anterior, posterior and inferior horns of the left lateral ventricles were manually traced separately in contiguous coronal brain sections, following detailed criteria with established inter- and intra-rater reliability [5,15].

2.3. Viscous Fluid Registration

We followed the approach proposed in [16] to deform the unlabeled image onto a target image on which the lateral ventricle had been manually traced. The deforming image was treated as a viscous fluid with motion governed by the Navier-Stokes equation, as pioneered by Christensen et al. [17]. This equation can be written as

$$\mu\nabla^2\vec{v} + (\mu+\lambda)\nabla(\nabla^{\mathrm{T}}\bullet\vec{v}) + F(\vec{x},\vec{u}) = 0\,, \qquad (1)$$

with $F(\vec{x},\vec{u})$ a force field which varies with the position $\vec{x}$ and drives the deformation $\vec{u}$, and with $\vec{v}(\vec{x},t)$ the deformation velocity experienced by a particle at position $\vec{x}$:

$$\vec{v}(\vec{x},t) = \frac{d\vec{u}(\vec{x},t)}{dt} = \frac{\partial\vec{u}(\vec{x},t)}{\partial t} + \vec{v}(\vec{x},t)\bullet\nabla\vec{u} \qquad (2)$$

The 3D convolution filter derived in [16] was used to implement a Green's function solution of the fluid equation. The summed squared intensity difference was chosen as the cost function, which is reasonable given the high contrast of ventricular CSF (though other cost functions are clearly possible [18]). Transformations resulting from the fluid registration were applied to the manually traced ventricular boundary using tri-linear interpolation, and generated a propagated contour on the unlabeled template images. **Fig. 1** shows an example of registration-based label propagation.

2.4. Ventricular Shape Modeling and Distance Mapping

To identify regional differences in ventricular morphology, we used surface-based mesh modeling to match homologous ventricular surface points across individuals [1,19]. Briefly, sets of points representing the tissue boundaries from each region were resampled and made spatially uniform by stretching a regular parametric grid (100 x 150 surface points) over each surface. Homologous grid-points from corresponding surfaces are then matched across subjects to obtain group average parametric meshes. For each surface model, a medial curve was defined as the line traced out by the centroid of the ventricular boundary ([1]; see [4] for related work on *M-reps*).
The local radial size was defined as the radial distance between a boundary point and its associated medial curve.

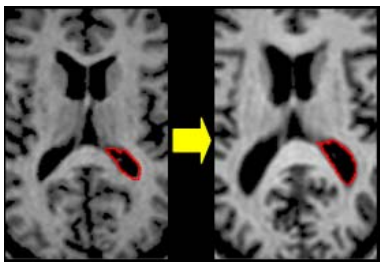

Fig. 1. Example of fluid registration-based surface propagation: the image on the left is the target image with a manually traced ventricular contour (*shown in red*). Through the fluid registration, the contour is mapped into the unlabeled image (*right panel*), matching the geometry of the posterior horn.

2.5. Ventricular Statistical Maps and Analysis

Surface contractions and expansions were then statistically compared between groups at equivalent locations [1,3]. Local group differences in ventricular shape between patients and controls were detected using Student's *t* tests at equivalent surface locations in 3D. The associated *P*-values describing the uncorrected significance of group differences were plotted onto the average surface model, producing a color-coded map.

2.6. Validation Study Design

To test our approach, surface models obtained automatically were compared to manual segmentations for the three horns of the lateral ventricles. The labeling error was measured using the symmetrized Hausdorff distance, as follows:
Let us first define the distance $d(p,S')$ between a point p belonging to a surface S and a surface S' as:

$$d(p,S') = \min_{p'\in S'} \| p - p' \|_2\,, \qquad (3)$$

where $\|.\|_2$ is the usual Euclidean norm. The Hausdorff distance $d(S,S')$ between S and S' is then given by:

$$d(S,S') = \max_{p\in S} d(p,S')\,, \qquad (4)$$

This distance is not symmetrical, i.e. $d(S,S') \neq d(S',S)$. We will refer to $d(S,S')$ as forward distance, and to $d(S',S)$ as backward distance. It is then convenient to introduce the symmetric Hausdorff distance $d_s(S,S')$ as follows:

$$d_s(S,S') = \frac{1}{2}[d(S,S') + d(S',S)] \qquad (5)$$

This quantity more accurately measures the error between two surfaces; a distance computed relative to a single surface can underestimate the true error.

3. RESULTS

3.1. Integrating Multiple Propagated Labels into One

Fig. 2a shows an example of a lateral ventricular surface extracted using different atlas images. If the registrations were perfect and there were no digitization errors, these images would all look identical. The right panel shows the average surface. The color bar represents the radial distance in mm.

By integrating multiple labels, effects of random digitization errors in each hand-traced segmentation are significantly reduced. The resulting average model is also somewhat robust to inaccuracies in individual registrations, which may occur when non-global minima of the intensity-based cost function are reached.

3.2. Comparison with Manual Labeling

Automatically propagated ventricular surface labels agreed closely with expertly labeled gold standard segmentations (**Fig. 2b)** in tests on a randomly selected set of 8 patients and 10 controls. Increasing the number of labeled atlases *N* resulted in an asymptotic decrease in the averaged symmetric Hausdorff error between manually and automatically extracted models.

3.3. 3D Ventricular Maps and Statistical Analysis

Statistical maps of mean radial distance comparing groups and maps of *P*-values for localized group differences are shown in **Fig. 2c,d.** As expected, patients with AD show localized volume enlargements in the posterior horn of the ventricles and in the mid-portion of the anterior horn (see [2]). The group-averaged ventricular surface models were automatically generated using 4 randomly selected atlases from the control group. *P*-value maps describing the significance of group differences were plotted onto the model surface. As shown in **Fig. 2d,** the total area of suprathreshold statistics (*red*) increased with *N*.

Under the null hypothesis of no group difference, a plot of the cumulative distribution function (CDF) of the observed versus null *P*-values is approximately diagonal. However, if significant changes are detected, the CDF increases rapidly from 0 over the small interval $[0, \alpha]$ and then levels off close to 1. Here we obtained *P*-values for a set of 60,000 surface points covering the anterior and posterior horns. Inferior horn points were hard to delineate manually and were not included, due to concerns that a manual 'gold standard' for that region might be error-prone itself. **Fig. 2e** shows the cumulative *P*-value distribution, and the effect of varying *N* from 1 to 4. The pink dashed line represents the expected CDF under the null hypothesis. As shown, the statistical power increases with *N*.

To determine the optimal value of *N*, we performed 2-tailed *t* tests to see by how much Hausdorff errors fell when adding an additional atlas, i.e., *t-test*(Error(*N*),Error(*N*+1)), for the anterior and posterior horns. In this study, values of N>3 did not detectably increase the power (**Fig. 2f**). Values shown above the points represent the associated *P*-values.

4. DISCUSSION

Compared with the 'single-atlas' approach, 'multi-atlas' fluid registration increased the statistical power in our surface-based morphometric study. The approach should generalize to deformable surface maps of the hippocampus and caudate; the optimal number of atlases to use depends on the distribution of registration errors for each structure, and whether there is any systematic bias between automated and manual segmentations. Models may also be better combined using approaches more robust to outliers, such as median filters on binarized labels. The accuracy and power gained may also depend on imaging protocol differences between the labeled and test images. The incremental value of this method versus any standard approach depends on the statistical power required. It may benefit drug trials where power is low and image acquisition is costly. This technique may also empower neuroscientific research to find specific genes, risk factors or neuroprotective treatments that modify rates of brain degeneration.

5. REFERENCES

[1] Thompson PM et al. *NeuroImage* 22:1754-1766 (2004).
[2] Carmichael OT et al. *IEEE Int. Symp. Biomed. Imag.* (2006).
[3] Thompson PM et al. *Neuroimage* 31(1):12-23 (2006).
[4] Styner M et al., *PNAS* 102(12):4872-4877 (2005).
[5] Narr KL et al. *Biol. Psychiatry* 50:84-97 (2001).
[6] Wang YL et al. *Proc. ICCV2005*, 1061-1066 (2005).
[7] Collins DL et al. *J Comput Assist Tomogr*. 18(2):192-205 (1994).
[8] Dawant BM et al. *IEEE-TMI* 18:909-916 (1999).
[9] Svarer C et al. *NeuroImage* 24:969-979 (2005).
[10] Heckemann RA et al. *NeuroImage* 32:70-78 (2006).
[11] Rohlfing T et al., *IPMI* 210-221 (2003).
[12] Rohlfing T et al., *NeuroImage* 21: 1428-1442 (2004).
[13] Klein A et al., *BMC Medical Imaging* 5:7 (2005).
[14] Kochunov P et al., *Hum. Br. Mapp.* 24:325-331 (2005).
[15] Woods RP. *NeuroImage* 19, 1829-1834 (2003).
[16] Gramkow C. *M.Sc. Thesis*, Danish Tech. Univ., (1996)
[17] Christensen GE et al., *IEEE-TIP* 5:1435-1447 (1996).
[18] Chiang MC et al. *NeuroImage* (2006).
[19] Thompson PM et al. *NeuroImage* 3:19-34 (1996).

Grant Support: NIH AG016570, LM05639, EB01651, RR019771 (to P.T.), P41 RR013642 (to A.W.T.).

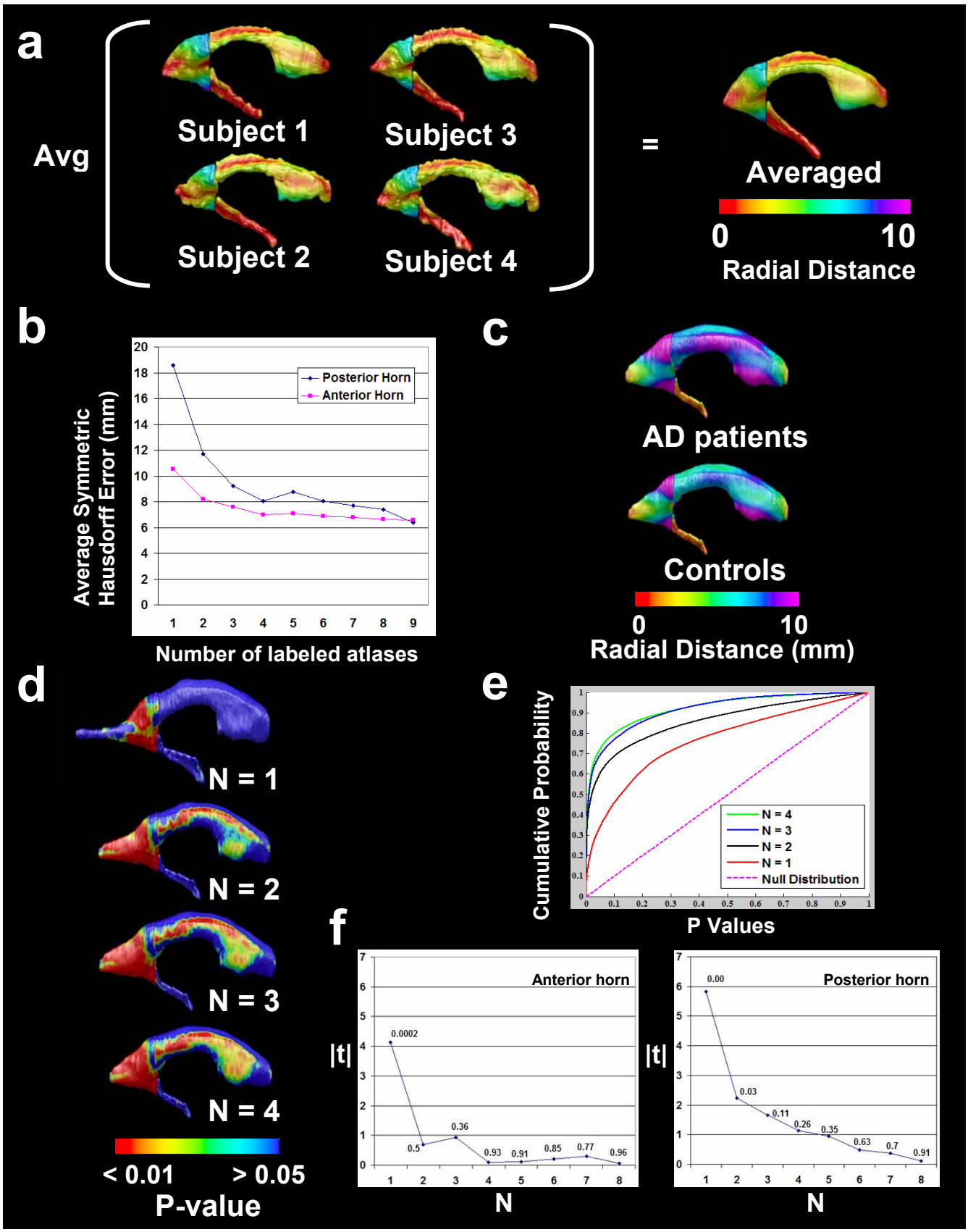
a
Avg
Subject 1
Subject 3
Subject 2
Subject 4
=
Averaged
0
10
Radial Distance
b
Average Symmetric Hausdorff Error (mm)
Posterior Horn
Anterior Horn
Number of labeled atlases
c
AD patients
Controls
0
10
Radial Distance (mm)
d
N = 1
N = 2
N = 3
N = 4
< 0.01
> 0.05
P-value
e
Cumulative Probability
N = 4
N = 3
N = 2
N = 1
Null Distribution
P Values
f
Anterior horn
|t|
0.0002
0.5
0.36
0.93
0.91
0.85
0.77
0.96
N
Posterior horn
|t|
0.00
0.03
0.11
0.26
0.35
0.63
0.7
0.91
N

Fig. 2.