Atlas-Based Hippocampus Segmentation In Alzheimer's Disease and Mild Cognitive Impairment

O. T. Carmichael [1], H. J. Aizenstein [2], S. W. Davis [1], J. T. Becker [3], P. M. Thompson [4], C. C. Meltzer [1], Y. Liu [5]

[1]Radiology, University of Pittsburgh, Pittsburgh, PA, United States, [2]Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States, [3]Neurology, University of Pittsburgh, Pittsburgh, PA, United States, [4]Neurology, University of California, Los Angeles, Los Angeles, CA, United States, [5]The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, United States

Introduction. Hippocampal atrophy is known to occur early in the course of Alzheimer's Disease (AD) on a spatial scale large enough to be detectable with structural MR images [1]. Visual, qualitative atrophy assessment has been hindered by the presence of low-contrast boundaries between neighboring anatomical structures, varying protocols for atrophy assessment, and the relative subtlety of atrophy early in the course of AD. Furthermore, manual delineation, or *segmentation*, of the hippocampus is labor-intensive, varies from person to person, and requires training the rater. For these reasons, manual segmentation of large numbers of hippocampi for broad studies of AD-related atrophy effects has not been feasible. Therefore, we investigate the feasibility of several fully-automated *atlas-based* methods for hippocampus segmentation in AD subjects, subjects with mild cognitive impairment (MCI) [7], and elderly controls. Atlas-based methods estimate hippocampi on subject images by geometrically aligning the subject image to a reference image (the atlas image) on which the hippocampus has been manually segmented (Figure 1).

Methods. Spoiled gradient-recalled (SPGR) volumetric T1-weighted MR images were acquired in the coronal plane for a set of 20,19, and 15 subjects in the AD, MCI, and control categories respectively at time of enrollment in the University of Pittsburgh Alzheimer's Disease Research Center between 1999 and 2004. Acquisition parameters were optimized for maximal contrast among gray matter, white matter, and CSF (TE=5, TR =25, flip angle = 40 degrees, NEX = 1, slice thickness = 1.5mm/0mm interslice). Subjects received a comprehensive battery of neuropsychological and clinical tests at time of enrollment [6], and a consensus meeting of neuroradiologists, psychiatrists, neurologists, and psychologists diagnosed each subject into MCI, AD, or control categories. Hippocampi were manually traced on all subject images by an expert rater who was highly consistent with two other trained raters (see [8] for tracing protocol). Widely-disseminated registration software packages (AIR [10], SPM [3], FSL [4], and Chen's method [2]) aligned the subject images to the standard MNI and Harvard atlas images [5] [9], and subject hippocampi were estimated by transferring traced hippocampi from the atlases to the subject image space. Additionally, subject hippocampi were estimated by aligning subject images to *cohort "atlases"*—that is, randomly-selected subject images augmented with manual tracings. Agreement between automatically-segmented hippocampi and the gold-standard manual tracings was assessed by quantitative measures of the degree of overlap between them. The fixed effects of registration method, side of the brain, disease state, choice of atlas, and manual tracing protocol, along with the random effects of subject selection and cohort atlas selection, were analyzed in mixed-effects statistical models with manual-automated segmentation overlap as the dependent variable. Registration methods were grouped into categories based on the degree to which they geometrically deform the images: affine methods rotate, translate and shear images at a global scale; semi-deformable methods locally deform images in a spatially smooth way; and fully-deformable methods deform images arbitrarily. Statistical analysis was performed using R verson 1.9.1.

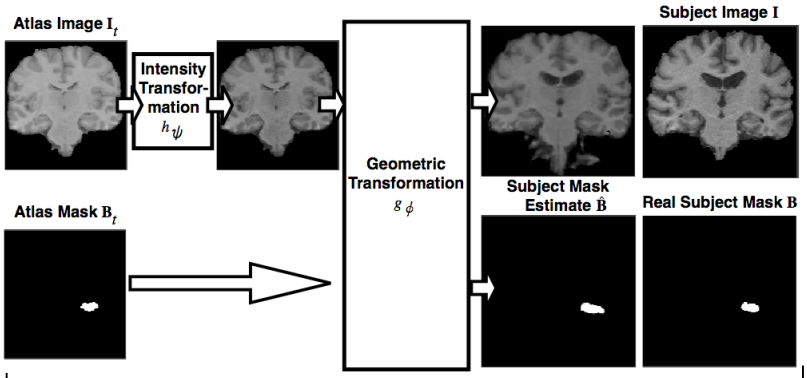

Figure 1. Schematic of the automated hippocampus segmentation technique.

Results. (Figure 2). The overall effects of side of the brain, disease state, and registration method on manual-automated overlap were statistically significant ($p<.0001$, $p=.0192$,$p<.0001$). Manual-automated overlap was significantly lower in AD compared to MCI ($p=0.0239$) and control ($p=.011$) groups. No significant difference was seen between MCI and control groups ($p=.647$). The methods, ranked in decreasing order of overlap ratio, were as follows: Chen fully-deformable, AIR semi-deformable, Chen semi-deformable, SPM affine, SPM semi-deformable, FLIRT affine, AIR affine. No significant difference existed between the FLIRT affine and AIR affine methods ($p=.286$), but for all other pairs of methods, the differences in overlap were significant ($p<.001$). Fully-deformable methods had significantly higher manual-automated overlap than semi-deformable and affine methods ($p<.001$), and semi-deformable methods had significantly higher manual-automated overlap than affine methods ($p<.001$). Using the Chen fully-deformable registration method, manual-automated overlap was significantly higher for automated segmentations based on alignment to a randomly-selected cohort atlas, compared to segmentations based on alignment to the MNI ($p<.001$) or Harvard ($p<.003$) atlas images. On a set of 2 AD, 2 MCI, and 2 control images manually traced by 3 human raters, differences between manual-automated overlap using the Chen fully-deformable method and overlap between pairs of human raters was not statistically significant ($p=.0916$).

Discussion. Fully-deformable image alignment methods are better suited for alignment-based automated hippocampus segmentation in elderly brains due to their ability to match the complex shapes of brain structures with high accuracy. Since atlas-based segmentation based on fully-deformable registration is competitive with manual segmentation in terms of manual-automated agreement, it is a promising approach for large-scale automated studies. Furthermore, extreme morphological differences between atrophied elderly subject brains and young, healthy controls may cause registration difficulties that make alignment to a randomly-selected cohort atlas image more accurate than alignment to the young-subject-based standard atlas images from MNI and Harvard.

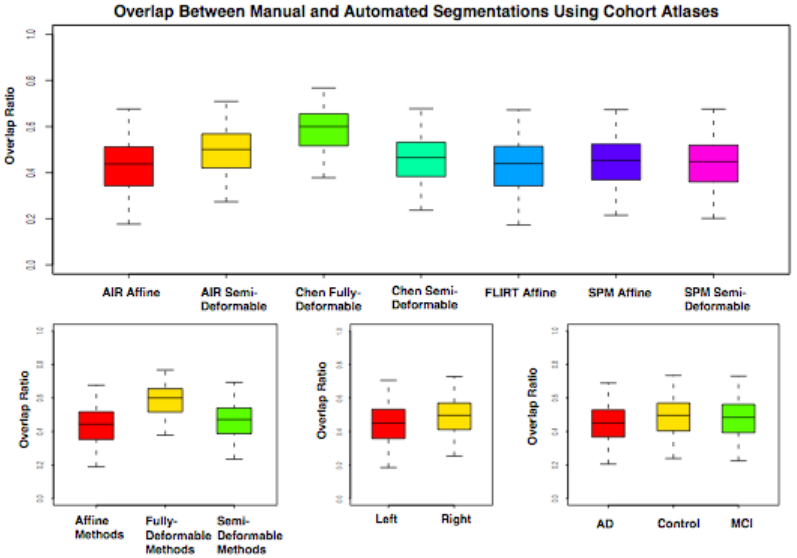

Figure 2. Performance of automated hippocampus segmentation.

References. [1] M Bobinski et al. Neurobio. Aging, 17(6):909–919, 1996. [2] M. Chen. PhD thesis, Carnegie Mellon University, October 1999. [3] K.J. Friston et al. Human Brain Mapping Ann. Mtg., pg.165–189, 1995. [4] M. Jenkinson et al. Technical Report, Oxford Centre for fMRI of the Brain, 2002. [5] R. Kikinis et al. IEEE Trans. Vis. Comp. Graph. 2(3):232–241, 1996. [6] OL Lopez et al, Neurology, 55:1854–1869, 2000. [7] RC Petersen et al, Int Psychogeriatr, 9 (suppl. 1):65–69, 1997. [8] P.M. Thompson, et al. NeuroImage, June 2004. [9] N Tzourio-Mazoyer et al, NeuroImage, 15:273–289, 2002. [10] RP Woods et al, J. Comp. Assist. Tomo, 22:139–152, 1998.