Quality of brain structure volume measurements using automatic anatomical segmentation based on propagation of manually defined labels

R. A. Heckemann[1], J. V. Hajnal[2], D. Rueckert[3], D. L. Hill[4], A. Hammers[5]

[1]Imaging Sciences Department, Imperial College London, Ascot, Berkshire, United Kingdom, [2]Imaging Sciences Department, MRC Clinical Sciences Centre, Hammersmith Hospital, Imperial College London, London, United Kingdom, [3]Department of Computing, Imperial College, London, United Kingdom, [4]Division of Imaging Sciences, King's College London, London, United Kingdom, [5]Faculty of Medicine, Imperial College at Hammersmith Hospital Campus, Neurology Group, MRC Clinical Sciences Centre, London, United Kingdom

Background

Automating the process of anatomical segmentation of brain MRI data sets has potential benefits: through analyzing large cohorts of normal brain images for the size, shape and location of individual structures, authoritative statistical reference sets could be generated. Similarly, patterns of abnormality could be studied in cohorts of patients with brain diseases. Currently, the preferred way to obtain reliable, accurate anatomical segmentations is through manual or semi-automated processes, which frequently require many hours of expert time for each subject [eg, 1]. Label propagation is a promising approach to automating segmentation, but it depends on how well the transformation applied to the labels represents anatomical correspondence between the atlas and the target data sets. In this work-in-progress, we investigate high-dimensional non-rigid registration [2] as a means to establish anatomical correspondence between a pair of brain volumes of individual subjects. Initial studies of this process have indicated that both accuracy and precision can be highly variable for different brain structures [3]. Using high-quality manual segmentations on 20 normal volunteers as the "gold standard" [1], we sought to establish the accuracy of size measurements and identify modes of failure of label propagation based on non-rigid registration.

Material and Method

We studied three-dimensional T1-weighted MRI brain images of 20 normal subjects. The data sets consisted of isotropic voxels with a volume of 0.823 mm^3. For each subject, an expert segmentation was available in a corresponding volume containing anatomical labels for 67 structures [1]. Each subject's labels were propagated onto each other subject by calculating a free-form deformation that maximizes normalized mutual information between the T1-volume pair [2], and applying this deformation individually to each label in isolation using trilinear interpolation. Structure sizes in the original and propagated label sets were determined as voxel counts. Interpolated voxels were counted according to their partial contribution to the structure size. Thus, for each structure in each subject label set, a manually determined reference size *s* as well as 19 label-propagation-based size estimates *l* were available. We used the percent relative deviation $rd = 100 \cdot (l-s)/s$ as the basic accuracy measure and looked for patterns and their possible causes.

Results

Label propagation predicted the size of the structure within 20% of the native segmentation in 70% of all measurements. Substantial variation was seen depending on the characteristics of the structure measured. In general, high boundary contrast and homogeneity within the structure appeared to facilitate more accurate measurements, as seen for example in the brain stem ($rd = 2.35 \pm 4.80$) and the left and right insula ($rd = 1.85 \pm 7.54$; 1.96 ± 7.14). Results for the temporal horns of the lateral ventricles and the nuclei accumbentes were particularly poor (left temporal horn *mean rd* = 1.60 ± 35.75, right accumbens *mean rd* = 4.49 ± 39.43), probably owing to the small size of these structures (average *s* for left temporal horn: 680 ± 179 voxels, right accumbens: 376 ± 81 voxels), the absence of a clear intensity boundary between nucleus accumbens and remainder of the striatum, partial volume effects and apparent discontinuity of the temporal horn. For the remaining structures, which ranged between 1253 and 86299 voxels (mean *s* across all subjects), structure size did not affect measurement accuracy. A commonality of many structures that were not estimated accurately was that their boundaries were partially defined as planes through landmarks, as opposed to grey-scale intensity gradients (Fig. 1). Such structures tended to be misestimated according to the difference between *s* in the subject pair (Δs correlated with *rd*), and the majority of extreme *rd* values were seen in this group (eg, left gyrus cinguli, anterior part: Δs to *rd* correlation 0.67, *rd* in one subject pair 252%).

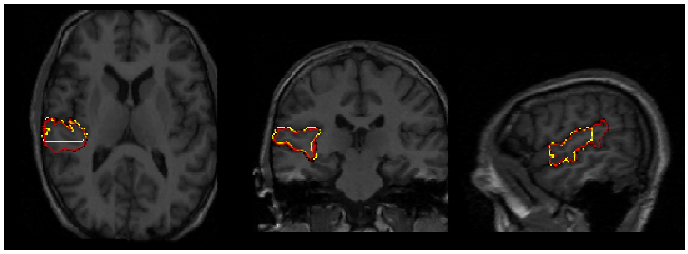

Fig. 1. Anterior part of the superior temporal gyrus. The outline of the manually generated label is shown in yellow, the outline of the propagated label in red. The posterior boundary is defined as a coronal plane through a landmark. This is misplaced by the label propagation procedure, resulting in a size overestimation of 47% in this particular case. Other, grey-scale gradient-based boundaries are closely matched.

Discussion and Conclusions

Registration-based label propagation as presented in this work facilitates automated anatomical segmentation and structure size measurements, especially of structures that are bounded by grey-scale gradients. Some structure boundaries were defined as coronal, transverse or sagittal planes that intersect with certain landmarks. Since the registration procedure does not take into account any knowledge about landmarks, such boundaries propagate with larger amounts of error.

The fact that large differences between structure sizes in target/source subject pairs reduce the accuracy of the label propagation suggests that an averaged MRI brain atlas, to be constructed from aggregated registration information, could further increase the accuracy of this approach.

Ongoing work will focus on exploring how to choose label definitions appropriately for segmentation propagation, tuning the registration procedure to improve intersubject anatomical correspondence, and using other measures to verify this approach to automated anatomical segmentation.

References

1. A Hammers, R Allom, MJ Koepp SL Free, R Myers, L Lemieux, TN Mitchell, DJ Brooks, JS Duncan. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Human Brain Mapping 19:224-247 (2003)
2. D Rueckert, LI Sonoda, C Hayes, DL Hill, MO Leach, DJ Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 18:712-721 (1999)
3. WR Crum, D Rueckert, M Jenkinson, D Kennedy, SM Smith. A Framework For Detailed Objective Comparison of Non-Rigid Registration Algorithms in Neuroimaging. In: C Barillot, DR Haynor, P Hellier (Eds.): MICCAI 2004, LNCS 3216, 679-686. Berlin, Heidelberg 2004.