Unbiased Robust Template Estimation for Longitudinal Analysis in FreeSurfer

Martin Reuter, H. Diana Rosas, Bruce Fischl
mreuter@nmr.mgh.harvard.edu

MGH/HST Athinoula A. Martinos Center for Biomedical Imaging

MASSACHUSETTS GENERAL HOSPITAL

Harvard-MIT Health Sciences & Technology

CSAIL

1 Background

Compared with cross-sectional studies, a **longitudinal design** can significantly reduce the confounding effect of inter-individual morphological variability by using each subject as his or her own control. As a result, longitudinal imaging studies are increasing in popularity in various aspects of neuroscience. Changes in gray matter that makes up the cortical sheet are for example manifested in aging, Alzheimer's disease, Huntington's disease, multiple sclerosis and schizophrenia. In vivo cortical thickness measures could be useful as **marker of disease progression or onset**; this is an active and important area of research. Longitudinal imaging-based biomarkers are thus of great potential utility in evaluating the efficiency of disease-modifying therapies. For these reasons, developing more robust and reliable measures of cortical, subcortical and white matter atrophy may have **a profound clinical impact**. The current methods that utilize cross-sectional approaches, in which images are processed individually introduce the natural variability of the brain as a confound. We sought to develop and validate a longitudinal approach that takes advantage of intra-subject longitudinal acquired scans to **improve the sensitivity and reliability** of automatic neuro-imaging morphometic measures.

2 Method

• The FreeSurfer cortical and subcortical segmentation and parcellation procedure involves **complex iterative nonlinear optimization** problems, such as topology correction, nonlinear atlas registration, and the nonlinear spherical surface registration. The final results can be **sensitive** to the selection of a particular starting point.

• Thus, by **initializing** the processing of time points in a longitudinal data series with the **same information**, we can reduce the variations and improve the robustness and sensitivity of the overall longitudinal analysis, making it possible to detect subtle longitudinal changes.

• In contrast to older methods (Han 2006) the longitudinal pipeline we present is designed to be **unbiased** with respect to any time point. Instead of initializing it with information from a specific time point (and thus creating a bias), a **template volume** is created and processed with FreeSurfer. It can be seen as an initial guess for the segmentation and surface reconstruction.

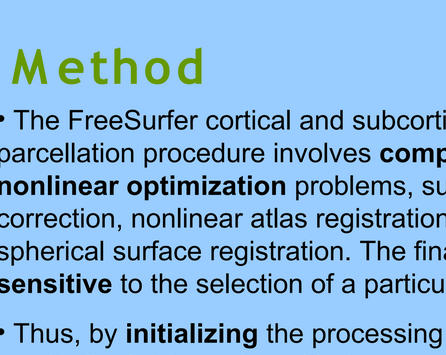

Figure 1: **Unbiased template estimation** for a subject with neurodegenerative disease: All time points are iteratively aligned to their **median image** with an inverse consistent robust registration method. Thus we obtain an unbiased robust template image for each subject and simultaneously a **co-registration of all time points**.

• Additionally we use **probabilistic voting** (temporal fusion) to consider label and intensity information across time at a specific location for the segmentation.

3 Results

As described above, the first step in the longitudinal analysis is the creation of an unbiased template together with the co-registration of all input images. Figure 2 shows the unbiased template of a series of 12 images (with atrophy) taken over a span of more than 4 years. The median is crisp as opposed to the more blurry mean.

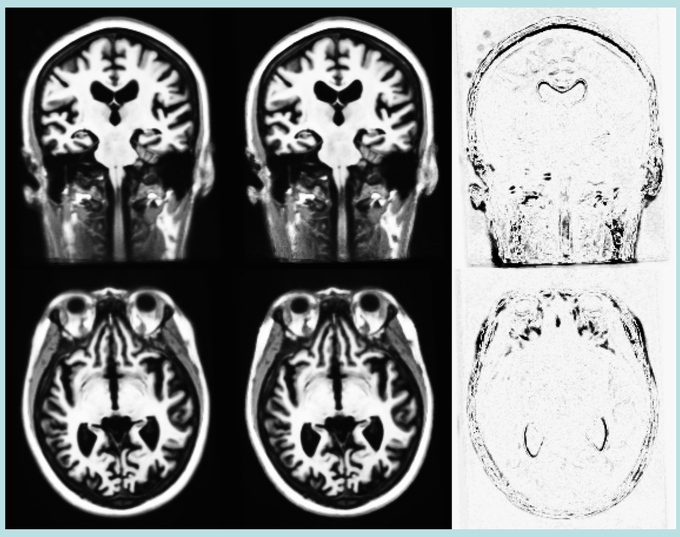

Figure 2: Comparison of **mean** (left) and **median** (middle) template image for a series of 12 images (4 years) of a subject with neurodegenerative disorder. The blur in the mean is visible in regions with that change over time (e.g. eye, neck). The difference image (right) between median and mean reveals significant blurring also at the ventricles (and corpus callosum, not shown). The average edge width is 5.93mm for the mean and 5.45mm for the median.

Test-Retest

In order to **evaluate the reliability** of the longitudinal scheme we analyze the variability in a test-retest study consisting of 14 healthy subjects with two time points (TP) taken 14 days apart. The images are T1-weighted MPRAGE full head scans (Siemens Sonata 1.5T). Figure 3 shows results of estimated subcortical volume change, comparing the independently processed time points (CROSS) and the longitudinal scheme (LONG). It can clearly be seen that LONG reduces variability in all regions, which leads to an increased power (or reduction of subjects) in a power analysis (Figure 4).

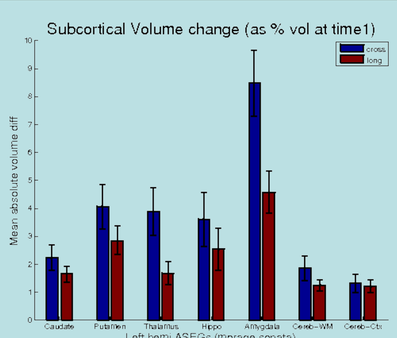

Figure 3**: Test-Retest** comparison of independent (cross) versus longitudinal (long) processing. The mean absolute volume difference (as % of volume at time 1) for several subcortical structures is shown. It can be seen that (long) significantly reduces variability.

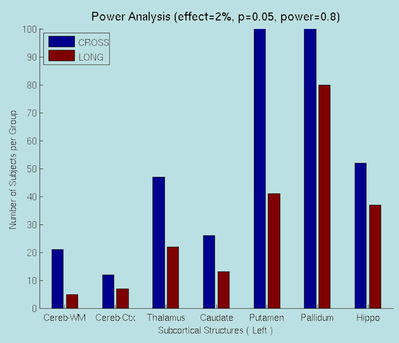

Figure 4: **Power Analysis**. Number of subjects needed to find a 2% volume loss at p=0.05 and 0.8 power. In the Thalamus, for example, the longitudinal stream needs only 20 subjects (10 in each group) as opposed to close to 50 when using independent (cross) processing.

Simulated Atrophy

To assess the sensitivity of the longitudinal analysis we applied approximately 2% simulated atrophy to the hippocampus in the left hemisphere of TP1 in the same population and used this synthetic image as TP2. As can be seen in Figure 5 (left) the longitudinal analysis not only detects the atrophy more accurately in the left hemisphere, but also correctly shows less variability and a zero mean in the right hemisphere (Figure 5, right) as opposed to the independent processing (cross).

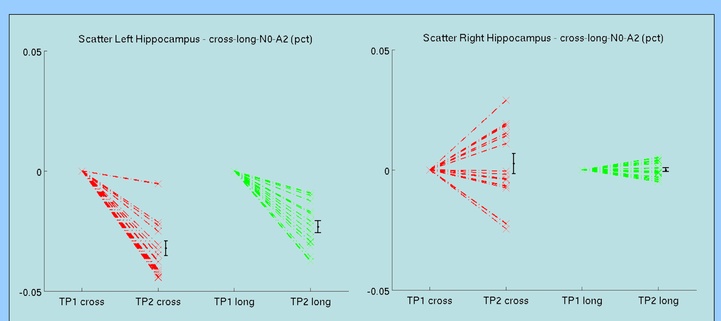

Figure 5: **Left hippocampus with 2% simulated atrophy.** By reducing the intensity of border voxels in the left Hippocampus, atrophy was simulated. The longitudinal processing (green) produces better results with smaller variance. **Right hippocampus without any change** to the image. The independently processed results (cross, red) show more variability, while the longitudinal results are highly accurate and correctly detect no change.

4 Conclusion

• We demonstrated that a **robust template image** can be taken as an **initial estimate** of the location of anatomical structures in a longitudinal scheme to improve accuracy and reduce variability of the automatically constructed segmentations in FreeSurfer.

• The scheme is completely **unbiased** with respect to any time point and therefore does not produce the common problem of differing results when switching or reversing the order.

• Due to the robust creation of the template, adding **additional** time points later is not expected to have a significant influence and therefore does not necessarily imply a re-processing of all time points if the template contains enough temporal information.

• The presented longitudinal scheme is freely available together with the software package **FreeSurfer** and has been successfully applied in our lab and by others for test-retest and longitudinal studies.

5 References

1. Han, X. & Jovichich, J., & Salat, D. & van der Kouwe, A. & Quinn, B. & Czanner, S. & Busa, E. & Pacheco, J. & Albert, M. & Killiany, R. & Maguire, P. & Rosas, D. & Makris, N. & Dale, A. & Dickerson, B. & Fischl, B. (2006), 'Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer', NeuroImage, vol. 32, pp 180-194.

Acknowledgements:

National Center for Research Resources P41-RR14075, BIRN002, U24 RR021382

National Institute of Neurological Disorders and Stroke R01 NS052585-01, R01 NS042861, P01 NS058793, NS058792

NIBIB R01 EB006758

National Institute on Aging, National Institutes of Health AG02238

Ellison Medical Foundation