SKELETON-BASED GYRI SULCI SEPARATION FOR IMPROVED ASSESSMENT OF CORTICAL THICKNESS

Mirco Richter [1], *Courtney A. Bishop* [2], *Juergen Dukart* [3,6], *Elisabeth Stuehler* [1], *Karsten Mueller* [3], *Matthias L. Schroeter* [3,4,5], *and Dorit Merhof* [1]

[1] Visual Computing, University of Konstanz, Germany
[2] FMRIB Centre, University of Oxford, UK
[3] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[4] Clinic of Cognitive Neurology, University of Leipzig, Germany
[5] Leipzig Research Center for Civilization Diseases, University of Leipzig, Germany
[6] LREN, Département des Neurosciences Cliniques, CHUV, Université de Lausanne, Lausanne, Switzerland

ABSTRACT

In order to improve classification of neurological diseases involving cortical thinning, this work proposes an approach for separating gyral and sulcal regions of the human cortex. Using data from magnetic resonance imaging, the skeleton of the brain's white matter was reconstructed and a geodesic distance measure was applied to separate gyri and sulci. Cortical thickness per subregion was measured for the entire cortex and for gyri and sulci individually in 21 patients with Alzheimer's disease, 10 patients with frontotemporal lobar degeneration composed of two subgroups and 13 control subjects. For discrimination using logistic regressions, which was assessed using leave-one-out cross-validation, improved results were obtained in five out of six group comparisons when cortical thickness measurements were constrained to gyral or sulcal regions.

Index Terms— Voxel-based gyri sulci separation, cortical thickness, dementia, magnetic resonance imaging.

1. INTRODUCTION

Neuropsychiatric diseases affecting the human cortex involve among others schizophrenia, mood disorders, autism, and dementia. To observe temporal changes per subject and to identify characteristic disease patterns in populations, valuable information is provided by structural measurements such as cortical thickness, gray matter volume or cortical folding patterns.

For classification in population studies, the datasets are usually spatially normalized to a common space either by voxel- or surface-based registration. In the case of high-resolution magnetic resonance imaging (MRI), both registration types offer the required accuracy to apply voxel- or vertexwise statistical analysis. Since these methods usually require a large amount of datasets to reduce the effect of noise, it is common practice to apply regional statistical analysis using mean values for predefined anatomical regions.

For this purpose, anatomical atlases are mapped to the anatomy of the subject's brain. Common registration methods comprise regions which include the entire cortex foldings. However, this might be insufficient due to differences in the cytoarchitecture of gyri and sulci. For this reason, a separate consideration of gyri and sulci is expected to improve statistical accuracy of cortical thickness analysis.

In this work, a voxel-based method for separating gyri and sulci and for assessing cortical thickness is presented and applied to datasets of dementia patients. As an advantage, voxel-based methods are generally faster than surface-based methods, and offer comparable accuracy.

2. MATERIAL AND METHODS

2.1. Subjects

The methods in this study were applied to structural MRI datasets of 13 control subjects (C), 21 patients with Alzheimer's disease (AD) and 10 patients with frontotemporal lobar degeneration (FTLD). The FTLD group comprised patients of two FTLD subtypes, six with semantic dementia (SD) and four with frontotemporal dementia (FTD). Probable AD was diagnosed according to the original and revised NINCDS-ADRDA criteria [1], FTLD according to the criteria suggested by Neary et al. [2]. The control group consisted of subjects with cognitive complaints that could not be verified in neuropsychological testing. The demographics are provided in Table 1 including scores of Mini Mental State Examination (MMSE) and Clinical Dementia Rating (CDR). MMSE scores for group C were not available.

The research protocol was approved by the ethics committee of the University of Leipzig, and was in accordance with the latest version of the Declaration of Helsinki. Informed consent was obtained from all subjects.

Category	C	AD	FTD	SD
Gender	6/7	12/9	1/3	3/3
Age	53.9±6.0	61.1±6.7	58.3±7.6	64.2±3.4
MMSE	n.a.	23.2±3.9	24.5±5.8	24.5±4.7
CDR	0.23±0.26	0.71±0.25	0.88±0.25	0.83±0.61

Table 1: Participant characteristics: number of female/male subjects, mean values and standard deviations of age (in years), MMSE and CDR scores.

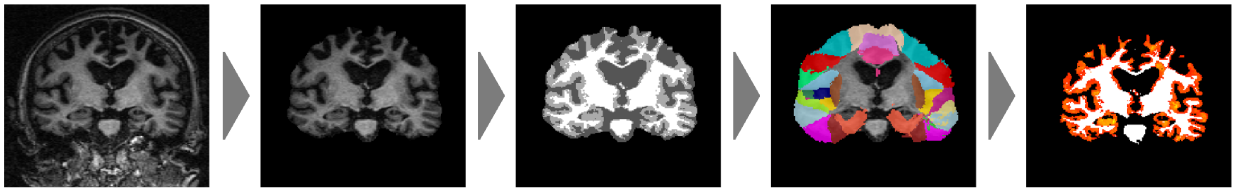

Fig. 1: Processing of MRI images.

2.2. Data acquisition

High-resolution T1-weighted MRI images were acquired on two different 3T scanners (MedSpec 30/100, Bruker Biospin, Ettlingen, Germany, and Magnetom Trio, Siemens, Erlangen, Germany) with two different sequences (MDEFT or MP-RAGE with TR =1300 ms, TI = 650 ms, TE = 3.93 ms or TE = 10 ms, FOV 25x25 cm^2, matrix = 256x256 voxels). Each scan comprised 128 sagittal slices adjusted to the AC-PC line, a slice thickness of 1.5 mm, and a pixel size of 1x1 mm^2. On the MedSpec scanner, only the MDEFT-sequence was used, whilst on the Magnetom Trio scanner, either MDEFT or MP-RAGE sequences were used. The distribution of scanner types and sequences used to obtain the MRI data was random across subjects and did not differ significantly in its distribution between the groups nor for scanner type nor for sequence.

2.3. Preprocessing

MRI images were preprocessed using the FMRIB Software Library (FSL, www.fmrib.ox.ac.uk/fsl) as depicted in Figure 1. At first, the skull was stripped with the Brain Extraction Tool [3] and the quality of each subject's brain mask was visually assessed for each slice by overlaying the mask on the original T1-weighted MRI. In case of missing tissue or included skull, the mask was manually corrected. Subsequently, the skullstripped brain tissue was segmented into white matter (WM), gray matter (GM) and cerebrospinal fluid by fitting a hidden Markov random field model including correction of the bias field [4]. The Gaussian mixture model provided the probabilistic voxel-wise membership values for each of the classes.

The common space for all subjects was provided by atlas registration to the MNI152 space by affine multi-resolution registrations using normalized correlation as cost function [5] and non-linear free-form deformations [6]. With the inverted warp fields, the labels of the Harvard-Oxford probabilistic atlas (distributed with FSL), which comprises 48 cortical regions of interest (ROIs) for each hemisphere, were transformed to the image space.

2.4. Gyri sulci cortical thickness estimation

The method for separating gyral and sulcal foldings proposed in this work is based on an approach for robust skeletonization of genus 0 objects in discrete volumetric images [7]. When applied to brain images, the WM is used for skeletonization since WM voxels of neighboring gyri do not touch (as opposed to GM voxels). Based on this skeletonization, a continuous function of the object's boundary is defined that is directly related to the curvature of the object's surface. While curvature-based surface classifiers are susceptible to noise, which is common for medical image segmentations, this approach provides a mechanism to detect noisy features on the object boundaries and to include them in the segmentation.

The skeleton approach is based on a distance map $F(p) \mapsto q_i$ that assigns to each point p inside the object Ω the closest points q_i on the boundary $\partial\Omega$ according to the Euclidean distance. Points p with $|F(p)| > 1$ are identified as skeleton points p_s, while an extension for discrete image data prevents holes in the skeleton. Then, for each skeleton point p_s the geodesic surface paths $\gamma(q_i, q_j)$ between all its associated boundary points are determined using a shortest path algorithm. The geodesic distance function $\rho(p_s)$ is then defined as the length of the longest path $\gamma(q_i, q_j)$, and a threshold τ_e is applied to separate gyri G' from sulci with $G'(p_s, \tau_e) = \{q_i = F(p_s) | \rho(p_s) < \tau_e\}$.

Since boundaries of GM segmentations are not smooth, the above separation would not only identify true gyri, but also detect noisy features (little bumps) inside sulci as gyri. Therefore, an additional geodesic distance function $\delta(q_i, R)$ is introduced in [7] which measures the distance to the remaining boundary points $R = \partial\Omega \setminus G'$. With a noise threshold τ_n, those boundary points with $\delta(q_i, R) < \tau_n/2$ are removed from G'. The final segmentation of gyri G without noisy features is then defined as

$$G(p_s, \tau_e, \tau_n) = \{q_i = F(p_s) | \rho(p_s) < \tau_e + \tau_n \wedge \delta(q_i, R) < \frac{\tau_n}{2}\}.$$

A threshold of $\tau_n = 10$ mm provides a good surface classifier for segmentations of brain images. The skeleton, associated boundary points and a geodesic path are depicted in Figure 2.

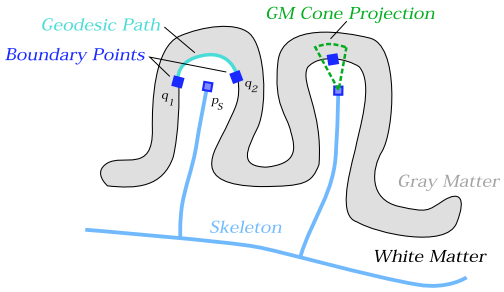

Fig. 2: Gyri sulci separation and projection of WM skeleton and associated boundary points to the GM medial layer.

Using the above separation of gyri and sulci, labels of the Harvard-Oxford probabilistic atlas are subdivided and the mean cortical thickness is measured for both gyri and sulci using minimum line integrals (MLI) with the parameters suggested in [8]. In order to measure cortical thickness robustly with the MLI approach, the sample points need to be restricted to the medial layer of the GM segmentation. For this purpose, the geometric relation between skeleton and boundary points is employed by extending their connecting line towards the GM layer and by determining the GM entry

and exit points as follows: As soon as a minimum number N_b of consecutive sample points S_{gm} classified as GM is detected, the entry point is identified as the first one of S_{gm}. Similarly, for the detection of the exit point, there must be N_b consecutive sample points without GM membership. N_b depends on the step width and the GM class membership is determined at each sample point using tri-linear interpolation of the GM segmentation.

After computing the GM entry and exit points, the point in the middle is defined as part of the GM medial layer where cortical thickness is measured with MLI. Due to the low resolution of the discrete image data and in favor of a smooth surface sampling, the projection to the GM medial layer is additionally applied for lines inside a cone with an opening angle of 5° that surrounds $\overline{p_s q_i}$ as illustrated in Figure 2.

2.5. Statistical analysis

To evaluate the gain in statistical power by applying the gyri/sulci separation for cortical thickness analysis, a discrimination model is applied. For this purpose, the cortical thickness data of each segmented brain is parcellated with seven different geodesic distance thresholds $\tau_e \in [8, 20]$ mm resulting in gyri and sulci separated at different depths. Each of these parcellations is overlayed on the ROIs defined by the atlas registration, resulting in the datasets G_{τ_e} and S_{τ_e}. Additionally, dataset W contains the mean cortical thickness of the combined gyri and sulci. For each of the newly created subregions, the mean cortical thickness is computed.

The discriminatory power is assessed for each dataset and each of the 96 regions with single variable logistic regression (LR) [9]. Subject age and gender are not included as covariates since no substantial differences were observed in previous studies [10]. According to the goodness-of-fit evaluated by the log-likelihood (LL), the best eight ROIs per dataset are selected and permuted to determine an optimal subset for a multivariate version of LR. The optimal combination C_R is found by maximizing the area under the receiver operator characteristic (AUROC) [11].

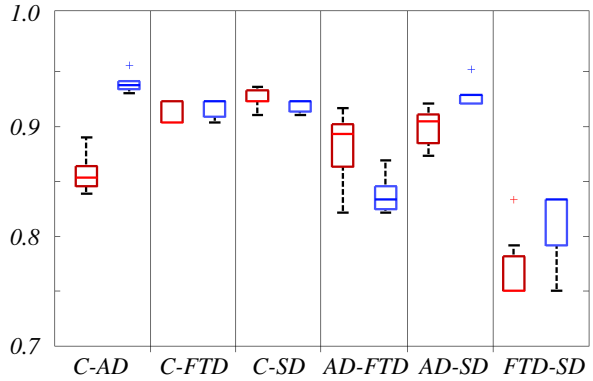

Fig. 4: Boxplots of AUROC values for each group pair for all gyri (red) and sulci (blue) datasets (G_{τ_e}/S_{τ_e}) indicating significant differences in discriminatory power for C-AD, AD-FTD and AD-SD (see Table 2).

3. RESULTS AND DISCUSSION

In Figure 3, the result of the gyri/sulci separation is illustrated by a projection on the outer cortical surface of the left hemisphere of one of the subjects. As compared to surface-based approaches such as Freesurfer [12], which took between 16 and 24 hours on a 3.07 GHz 8 core i7 Pentium CPU with 12 GB RAM for all processing steps, the performance of the presented approach amounted to 68 minutes for one sample, including preprocessing, gyri/sulci separation and cortical thickness measurement.

In order to assess whether different discrimination results are obtained when gyri or sulci are considered separately, Figure 4 shows boxplots of AUROC values of multivariate LR for G_{τ_e} and S_{τ_e} datasets of each group pair. For C-AD, AD-FTD and AD-SD, AUROC values differed significantly between G_{τ_e} and S_{τ_e} which was tested with non-parametric tests of Mann-Whitney at a significance level of $\alpha < 0.05$ ($H_0(G = S)$ in Table 2).

The optimal datasets for each group pair are summarized in Table 2. The discrimination results determined by leave-one-out cross-validation are provided, i.e. accuracy, sensitivity, specificity and AUROC along with the indices of the optimal combination of ROIs C_R in the multivariate LR. In five out of the six group pairs, discrimination on either a G_{τ_e} or a S_{τ_e} dataset performed better than on the combined dataset W according to AUROC (C-AD: 1.16%, C-FTD: 4.29%, C-SD: 4.35%, AD-FTD: 7.00%, AD-SD: 0.85%). In FTD-SD, the performances on datasets W and G_8 were equally good.

The best discriminating regions are stated in Table 2. Frontal operculum cortex left and planum polare right discriminated best for C-AD, Heschl's gyrus right for C-FTD, and inferior temporal gyrus left for C-SD. AD-FTD were separated with anterior cingulate gyrus left and posterior parahippocampal gyrus left, AD-SD with anterior middle temporal gyrus left, and FTD-SD with superior parietal lobule right.

Figure 5 highlights the LL values of single variable LR for the group pairs C-AD, C-FTD and C-SD. Similar results, both in discriminatory power and observed regions, were obtained in a study on cortical thickness measured with Freesurfer and a discriminatory approach with LR [10]. The identified regions correspond also to a meta-analysis of AD and FTLD with the exception of disease unspecific occipital regions [13]. Their dominance can be explained by the small group sizes.

According to the achieved accuracy and AUROC values, a clear discrimination between controls and Alzheimer's disease subjects, and between Alzheimer's disease and the two subtypes of frontotemporal lobar degeneration could be observed. Discrimination between the clinical groups, especially between FTD and SD suffered mainly from the small number of subjects.

4. CONCLUSION

In this work an approach was proposed for separating gyri and sulci on discrete neuromedical data to improve the discriminative power of the applied statistical model. The separation was accomplished by applying a robust skeletonization approach that defines a continuous pruning function using geodesic distances. With the help of cortical thickness data acquired for dementia patients, the discriminative power could be improved when gyri and sulci were considered separately in the logistic regression framework.

5. ACKNOWLEDGEMENTS

This work was supported by the DFG Research Training Group GK-1042 "Explorative Analysis and Visualization of Large Information Spaces" at the University of Konstanz.

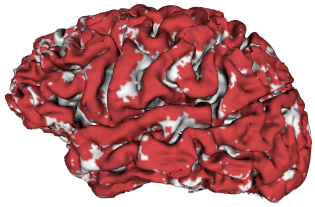

Fig. 3: Subdivision of gyri (red) and sulci (white) with $\tau_e = 8$ mm projected on the pial surface.

Group Pair	Dataset	Accuracy	Sensitivity	Specificity	AUROC	C_R	$H_0(G = S)$	$G \lesseqgtr S$
C-AD	S_{20}	85.3	85.7	84.6	95.6	41,92	5.83×10^{-4}	S
C-FTD	G_8	88.2	75.0	92.3	92.3	93	0.59	-
C-SD	G_{10}	78.9	66.7	84.6	93.6	14	0.39	-
AD-FTD	G_{12}	88.0	50.0	95.2	91.7	29,35	0.02	G
AD-SD	S_{14}	85.2	66.7	90.5	95.2	11	2.33×10^{-3}	S
FTD-SD	W	80.0	83.3	75.0	83.3	66	0.07	-

Table 2: Classification results of best discriminating datasets after multivariate LR for each group pair. Accuracy, sensitivity, specificity and AUROC are provided in %. C_R lists the ROI indices of the Harvard-Oxford atlas that discriminate best in multivariate LR. H(G=S): p-value of Mann-Whitney test for equal distribution of AUROC values for G_{τ_e} and S_{τ_e}. $G \lesseqgtr S$ states whether LR performed better on gyri (G) or sulci (S).

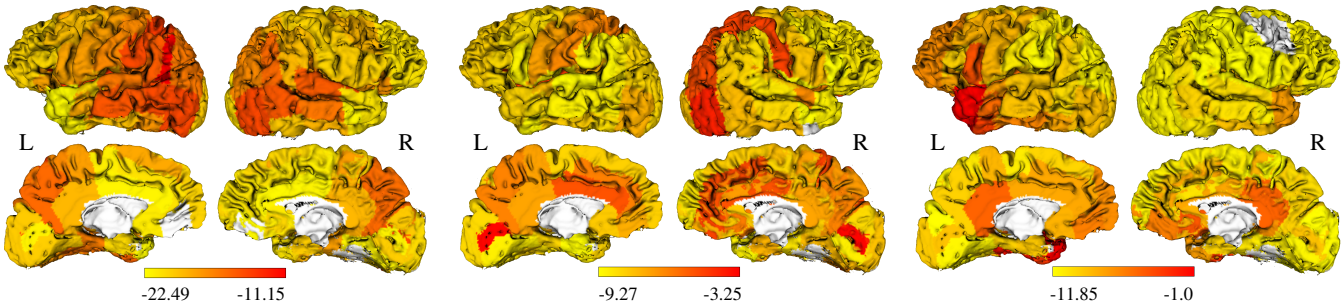

Fig. 5: ROIs highlighted according to log-likelihood of single variable logistic regression on pial surfaces of left and right hemisphere for groups $C\text{-}AD$ (left), $C\text{-}FTD$ (middle) and $C\text{-}SD$ (right). Red colors indicate regions of best, yellow and white of no model fit.

The authors J.D. and M.L.S. have been supported by LIFE - Leipzig Research Center for Civilization Diseases at the University of Leipzig. LIFE is funded by means of the European Union, by the European Regional Development Fund (ERFD) and by means of the Free State of Saxony within the framework of the excellence initiative. Furthermore, M.L.S. is supported by the German Federal Ministry of Education and Research (BMBF) as a member of the German FTLD consortium.

6. REFERENCES

[1] B. Dubois, H.H. Feldman, C. Jacova, S.T. Dekosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha, K. Meguro, J. O'brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, Y. Stern, P.J. Visser, and P. Scheltens, "Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria," *Lancet neurology*, vol. 6, no. 8, pp. 734–46, 2007.

[2] D. Neary, J.S. Snowden, L. Gustafson, U. Passant, D. Stuss, S. Black, M. Freedman, A. Kertesz, P.H. Robert, M. Albert, K. Boone, B.L. Miller, J. Cummings, and D.F. Benson, "Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria," *Neurology*, vol. 51, no. 6, pp. 1546–54, 1998.

[3] S.M. Smith, "Fast robust automated brain extraction," *Hum Brain Mapp*, vol. 17, no. 3, pp. 143–55, 2002.

[4] Y. Zhang, J.M. Brady, and S.M. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *Proc. IEEE Trans Med Imaging*, vol. 20, no. 1, pp. 45–57, 2001.

[5] M. Jenkinson and S.M. Smith, "A global optimisation method for robust affine registration of brain images," *Med Image Anal*, vol. 5, no. 2, pp. 143–56, 2001.

[6] J. Andersson, M. Jenkinson, and S. Smith, "Non-linear optimisation," technical report TR07JA1, FMRIB, 2007.

[7] D. Reniers, A. Jalba, and A. Telea, "Robust classification and analysis of anatomical surfaces using 3D skeletons," *Eurographics Workshop on Visual Computing for Biomedicine*, pp. 61–8, 2008.

[8] I. Aganj, G. Sapiro, N. Parikshak, S.K. Madsen, and P.M. Thompson, "Measurement of cortical thickness from MRI by minimum line integrals on soft-classified tissue," *Hum Brain Mapp*, vol. 30, no. 11, pp. 3188–99, 2009.

[9] D.W. Hosmer and S. Lemeshow, *Applied logistic regression*, John Wiley and Sons, 2000.

[10] A.-T. Du, N. Schuff, J.H. Kramer, H.J. Rosen, M.L. Gorno-Tempini, K. Rankin, B.L. Miller, and M.W. Weiner, "Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia," *Brain*, vol. 130, no. 4, pp. 1159–66, 2007.

[11] J.A. Hanley and B.J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[12] B. Fischl and A. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," *Proc Natl Acad Sci U S A*, vol. 97, no. 20, pp. 11050–5, 2000.

[13] M.L. Schroeter, T. Stein, N. Maslowski, and J. Neumann, "Combined imaging markers dissociate Alzheimer's disease and frontotemporal lobar degeneration - An ALE meta-analysis," *Front Aging Neurosci*, vol. 3, no. 10, 2011.