A Comparative Evaluation of Cortical Thickness Measurement Techniques

P. A. Bromiley*, M.L.J. Scott, and N.A. Thacker

Imaging Science and Biomedical Engineering, Stopford Building,
University of Manchester, Oxford Road, Manchester, M13 9PT.

Abstract. *In vivo* measurements of cortical thickness from MR images have potentially widespread utility in the characterisation of normal brain development and maturation as well as in diagnosing and measuring the progress of a number of cortical pathologies. The literature describes several approaches to this problem, which may be divided into two groups: those relying on deformable models of the inner and outer cortical surfaces, and those relying on image intensities alone. Results from the former may be largely model driven at points deep within the sulci, where no apparent channel of cerebrospinal fluid can be seen at the resolution of typical MR images, potentially introducing bias. We present a comparative evaluation of cortical thickness measurement techniques, which demonstrates that approaches based on edge detection can provide cortical thickness measurements of equal accuracy to model-based approaches, using less processor time, and without the possibility of bias from a model.

1 Introduction

The human cerebral cortex makes up the largest part of the brain, and consists of a highly convoluted layer of neuronal cells with the topology of a 2D sheet, surrounding a core of white matter. Its thickness varies considerably, from approximately 2mm in the calcarine sulcus to approximately 4mm in the precentral gyrus, with an average of approximately 3mm [1]. Measurements of cortical thickness have shown considerable potential both in the study of normal brain growth and maturation [2–6] and in the diagnosis, or measurement of the progression, of a wide variety of cortical pathologies including Alzheimer's disease [7], schizophrenia [7], and others [6]. Reliable and automated techniques for *in vivo* cortical thickness measurement therefore form a useful tool in neurology.

Any cortical thickness measurement technique requires two core components: a method for locating the inner and outer cortical surfaces, and a metric with which to measure the distance between them. A variety of thickness metrics have been suggested, varying from simply measuring the thickness along normals to the inner cortical surface [8] to approaches based on partial differential equations [9–11]. However, the thickness metric chosen must be considered the definition of the quantity to be measured: the accuracy with which this measurement is made is dictated by the method chosen to define the cortical surfaces. The techniques presented in the literature can be coarsely divided into two groups in this respect: model-based approaches that involve fitting deformable models to the inner and outer cortical surfaces i.e. the boundaries between white matter (WM) and grey matter (GM) and GM and pia matter (PM), and data-driven approaches that detect these interfaces using image intensities alone. The deformable model based approach typically involves segmenting the WM, fitting a model to the WM/GM interface, and then expanding this surface until it reaches the GM/PM interface. Many such algorithms include topological constraints that prevent self-intersections in the models [8, 12], thus ensuring that they have a simple, spherical topology, which has advantages in defining the distance metric and allows the results to be displayed as projections onto the cortical surface. This approach is used in the ASP (Anatomic Segmentation using Proximities) algorithm [8], which introduces both a surface self-proximity term and a term based on the distance between corresponding vertices on the inner and outer surface models.

The deformable model based approach has two main drawbacks. First, such algorithms require considerable computational resources, largely due to the topological constraints. For example, the algorithms presented in [8] and [12] required 30hrs on a 180 MHz Silicon Graphics R10000 and 5 hours on a 500 MHz Pentium 3 respectively to process each data set. Second, the introduction of terms to prevent self-intersections of the model surfaces may bias the algorithm towards a fixed separation between the inner and outer cortical surface models [8], depending on the weights assigned to each term. Such terms may be required in order to solve the notoriously difficult problem of fitting the model in tightly folded gyri [13, 14], where the sulcal banks oppose so closely that there is no clear CSF channel in the sulcus at the resolution of typical MR images. In such regions the outer surface model may fail to fit the pial surface within the sulcal fundus, leading to thickness estimates that are at least two times too high. Such biases may therefore be inevitable. These problems have led to an interest in data-driven techniques, in which the inner and outer cortical surfaces are determined using only image intensities (e.g. [11, 14, 15]). Such approaches also typically begin with segmentation of the WM, GM and cerebrospinal fluid (CSF). Thickness measurements are then performed at each

*E-mail: paul.bromiley@man.ac.uk

point on the inner cortical surface, by propagating away from the surface according to some thickness metric until the outer cortical surface is reached. One notable advantage of such approaches is that, in regions of tightly folded gyri as described above, if the pial surface is missed and another point on the inner cortical surface is reached, the thickness measurement can be halved to produce a value which is approximate but still entirely data-driven and so free from model bias. It has been shown that this approach has little effect on final, regional thickness measurements if local smoothing is applied [11, 14].

In previous work we have presented a data-driven cortical thickness measurement technique [16]. In this paper, we compare the results from this algorithm to a wide range of published measurements, produced using both model-based and data-driven algorithms. The aim was to determine whether the use of deformable models to define the inner and outer cortical boundaries increases the accuracy of the thickness measurement (i.e. reduces the random errors) or introduces bias (i.e. systematic errors).

2 Method

The cortical thickness measurement technique used in this study has been presented previously [16]; we summarise the method here. Several stages of preprocessing were used. Initially, the partial volume segmentation technique described in [17], which involves fitting a Bayesian mixture model containing both pure tissue and partial volume terms to the image histogram, was applied to the GM, WM and CSF. This process produced measurements of the means and standard deviations of the pure tissue grey levels and measurements of the most likely tissue volume contributions in each voxel. To obtain a finer through-plane resolution whilst preserving tissue boundaries, the data was explicitly up-interpolated using a partial volume scheme to constrain the potential tissue boundaries, determined using 3D image gradients, that could pass through a partial volume voxel. Finally, a map of the GM was produced. In addition, the original image data was registered to a common stereotaxic space (the Talairach atlas [18]) using a linear affine transform. The atlas defined the 31 cortical regions used later in producing regional histograms of the cortical thickness: its use allowed direct comparison to results presented in [1].

The cortical thickness was measured using a modified edge detection process (Canny [19]) to determine the GM/WM boundary. A 'z-score' measure of the grey-level of each voxel being consistent with the GM/WM midpoint value was used to construct a likelihood image which highlighted the GM/WM boundary. This was then used as a replacement for the conventionally used sum-squared image gradient (edge strength) map in the Canny edge detector in order to produce well localised connected edge strings to sub-pixel accuracy. The 3D surface normal at each voxel on the GM/WM boundary was determined by taking the local grey-level gradient of the 3D, Gaussian-smoothed (using a kernel of [1/2, 5]), up-interpolated grey-level data. The GM tissue probability maps and the GM/WM boundary and edge orientations were used to determine the distance from the boundary, at each voxel on the boundary, along the orientation direction to another GM edge. This edge could either be a GM/CSF interface or, if there was no CSF visible in the intervening sulcus, a GM/WM interface, and was determined by comparing the value of the voxel in the original grey-level image to the mean GM value. If it was a WM boundary it was assumed that the opposing banks of a sulcus had been traversed and the sulcal thickness was taken to be half this length: this approach has been shown to have little effect on subsequent regional thickness measurements [11, 14]. Histograms of the cortical thickness measurements in each region were then produced and their median values taken to produce the final regional thickness measurements. The regional histograms typically contained between 100 and 6000 entries, thus giving reasonable anatomical precision as well as a robust estimate of the median.

The method was applied to inversion-recovery MR scans (1.5T Philips ACS PT 6000 NT, TI/TR/TE = 300/6850/18 ms, pixel size = 0.9x0.9mm, 51 slices) of 119 normal volunteers (52 male, mean age = 70.3 years, range = 19-86 years). 110 scans had axial slices (thickness = 3.0mm), 9 had coronal (thickness = 4.0mm). The results were then compared to cortical thickness measurements from the literature in two stages. First, regional average thickness measurements in a range of standard cortical regions were computed and compared to the results published by Kabani et al. [1], which were produced using a deformable model-based technique. This analysis acted as an exemplar comparison between model-based and data-driven techniques, highlighting any difference in both random and systematic (i.e. model bias) errors. Second, a meta-study of precentral gyrus thickness variation with age was performed, involving measurements from nine studies including this one. This region is the location of the primary motor cortex, and is also the thickest region of the cortex; however, the choice of region was solely based on the number of previously published thickness measurements available in the literature. It is known that changes in both cortical thickness [10] and overall brain volume [20] occur with age. Any bias introduced by the use of deformable models would suppress the ability to observe such changes, and so this meta-study provides a more detailed evaluation of such biases.

3 Results

Kabani et al. [1] presented regional average cortical thickness measurements in ten regions from a group of 40 subjects, produced using the ASP algorithm described above. In addition, they presented manual measurements of the same quantities produced by expert observers. Figure 1a shows a comparison of these results to thickness measurements in the same regions produced using the algorithm presented here. Linear fits to the rate of thickness change with age were used to correct these results to the equivalent thicknesses at the mean age of the Kabani et al. subject group, allowing a direct comparison to be performed. Assuming that the manual measurements provide a gold standard, Fig. 1b provides a more quantitative comparison, showing the difference between the manual measurements and the results from each algorithm. The mean differences across all regions are 0.61 ± 0.43 mm for the ASP algorithm and -0.21 ± 0.22 mm for the algorithm presented here. Neither difference is statistically significant, and so neither algorithm shows evidence of systematic errors on the basis of these results. In addition, the ratio of the random errors on the two sets of results ($0.43/0.22 = 1.95$) is as expected given the number of subjects included in each study ($\sqrt{119/40} = 1.72$).

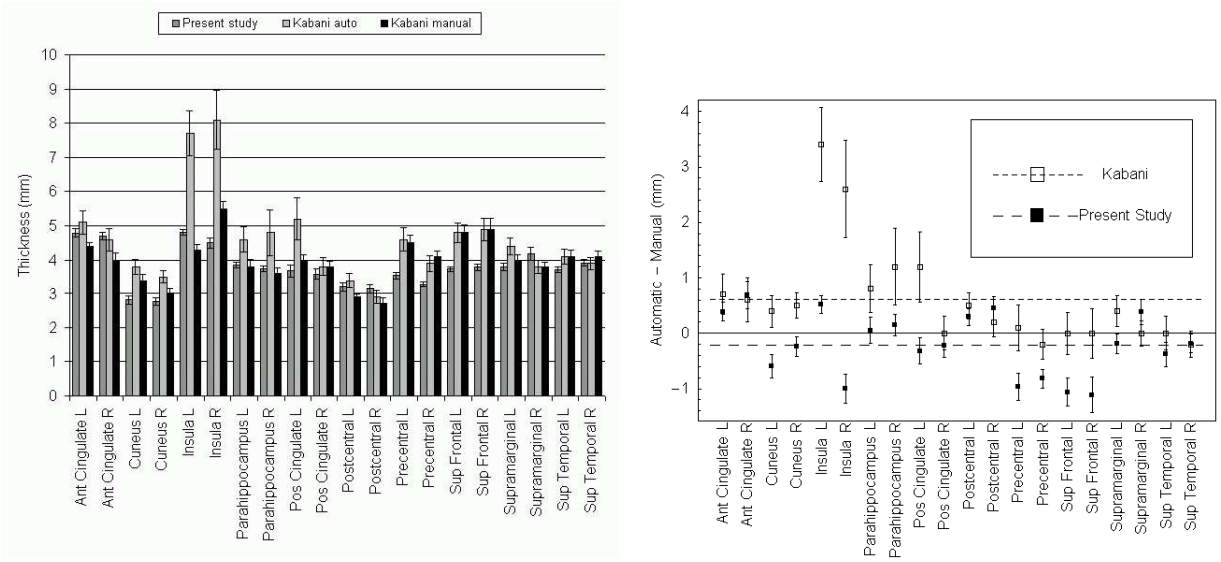
Reference	No. subjects	Age range (years)	Algorithm type
Kabani et al. 2001 [1]	40	18-40	Model based
von Economo 1929 [21]	-	30-40	Manual measurement
Sowell et al. 2004 [10]	45	5-11	Intensity based
Tosun et al. 2004 [22]	105	59-84	Model-based
Fishl et al. 2000 [12]	30	20-37	Model-based
Thompson et al. 2005 [14]	40	18-48	Intensity based
MacDonald et al. 2000 [8]	150	18-40	Model based
Salat et al. 2004 [5]	106	18-93	Model based

Table 1: Details of the studies included in the meta-study of the dependence of precentral gyrus thickness on age.

Figure 2 shows the average cortical thickness measurements for the precentral gyrus in 119 subjects produced using the algorithm presented here, plotted against age. A quadratic fit to the data is shown: the dashed curves either side of the fit show the upper and lower standard error bounds. A significant ($P < 0.0001$) reduction of cortical thickness with age is observed. In previous work [16] we found similar dependencies in other cortical regions. Also shown are a number of measurements of precentral gyrus thickness from the literature. The details of the studies involved are given in Table 1. With the exception of the results from Kabani et al. [1], the data were read from graphical representations[1]. The results presented by von Economo [21] were measured manually post-mortem: brain volume decreases by approximately 10% during postmortem fixation [23]. However, the thickness measurement was performed only on the gyral cap, which is known to be thicker than the sulcal fundi [8]. Similarly the presentation of the results in Sowell et al. [10] and Thompson et al. [14] as projections onto the outer cortical surface prevented identification of the thickness in the sulcal fundi. These three data therefore represent upper limits on the average thickness in the region. Overall, the studies represented in Fig. 2 represent the widest possible range of methods for defining the inner and outer cortical surfaces, the thickness metric, and the presentation of the results. Some variation between the measurements introduced by these differences in experimental procedure might therefore be expected. However, the data show a remarkable level of agreement. If the errors on the results quoted by Kabani et al. [1] can be taken as representative of the errors on the other studies, then there is no statistically significant difference between any of these results and our own.

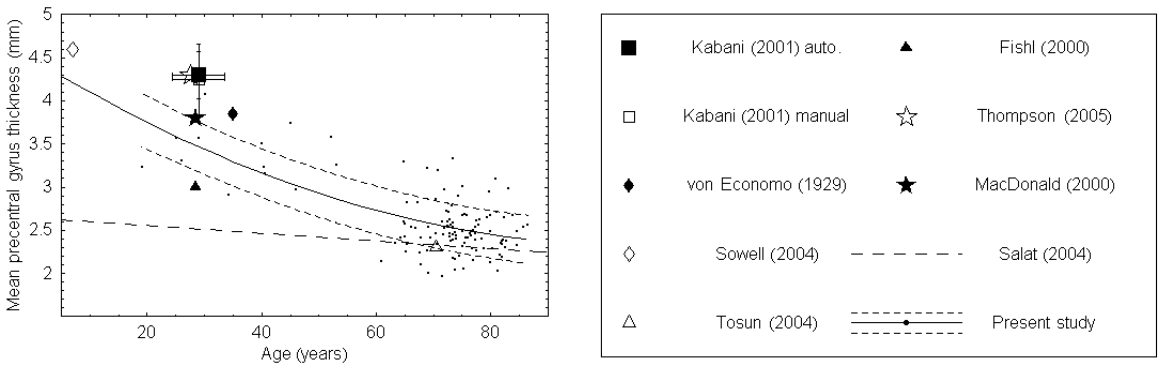
The remaining paper included in this study, Salat et al. [5], was the only one to study variation in cortical thickness with age, and the only model-based study to cover the whole age range between adolescence and senescence. A significant disagreement between this study and the others included in the meta-study can be seen, with Salat et al. suggesting a much lower rate of cortical thickness change with age. Given the level of agreement between the other studies, and the fact that the algorithm used by Salat et al. was based on deformable modelling of the inner and outer cortical surfaces, we suggest that this may be due to the possibility for bias identified by [8] in such algorithms. The constraints applied to prevent self-intersection of the models and to aid in the modelling of the surfaces in tightly folded gyri tend to bias the model towards a fixed thickness, suppressing the observation of age-related changes. The same effect was found in the mean rate of thickness change over the whole cortex: Salat et al. quote 0.0016 mm year^{-1} , whilst the present study gave a value of 0.016 ± 0.0052 mm year^{-1} i.e. an order of magnitude higher. For comparison, a global rate of $0.0077\ mm\ year^{-1}$ was found in [2], consistent with our own result but not with that of Salat et al. Finally, age-related

[1]With the exception the results from Salat et al. [5] and von Economo [21] these were presented as views of the outer cortical surface, in some cases partially inflated to reveal the sulcal fundi, with colour coding to represent the thickness at each point. This method of data display is popular in the literature as it avoids the need for parcellation of the data into particular regions. However, the calculation of regional average thicknesses from such representations is difficult and the calculation of errors on the averages impossible. Hence, these points are shown without error bars.



(a) (b)

Figure 1. Manual and automatic regional average cortical thickness measurements presented in [1], and automatic results from the algorithm presented here (a), and differences between the manual thickness measurements and the algorithm results (b); the dashed lines show average differences across all regions for each algorithm.



(a) (b)

Figure 2. Measurements of the average cortical thickness in the precentral gyrus: see main text for description.

changes in GM density i.e. the proportion of GM within a kernel around each point on the cortical surface, a quantity closely related to cortical thickness, were measured in [4]: the average, proportional rate of change of GM density in the precentral gyrus was $0.593\ \%\ year^{-1}$, consistent with the $0.567 \pm 0.270\ \%\ year^{-1}$ rate of change in cortical thickness found in the present study.

4 Conclusion

The literature describes several approaches to cortical thickness measurement, many of which rely on measuring the distance between deformable models fitted to the inner and outer cortical surfaces. The topological constraints applied in such algorithms result in a requirement for considerable amounts of processor time, typically several hours per image volume. In addition, bias may be introduced by the model at points where the information available from the image intensities is weak. In previous work we presented a simpler alternative, based on edge detection followed by measurement of the thickness along normals to the inner cortical surface, which requires an order of magnitude less processor time and does not introduce bias. In this paper we have presented comparisons of the results from this algorithm against a range of measurements from the literature. The comparisons indicate that results from this algorithm are no less accurate than those published in the literature, suggesting that the accuracy of both types of algorithm is dictated by the initial segmentation. However, in at least one study, the use of an algorithm based on deformable models appears to have suppressed the ability to detect age-related change. We therefore conclude that the use of deformable models in cortical thickness measurement provides no advantage in terms of accuracy.

The comparison of age-related changes in cortical thickness can also be used to investigate the ability of the algorithm presented here to detect pathological effects. Comparison with previous results shows that the thickness measurements in the precentral gyrus are consistent with those presented in the literature over a wide age range, confirming the accuracy of the technique. In previous work [16] we have shown that statistically significant age-related changes were detected in this data set throughout the cortex. Previous studies on age-related and pathological brain volume change [24] have shown that pathological changes typically occur an order of magnitude faster than age related changes. Therefore, the ability to detect age-related changes strongly implies that pathological changes will also be detected.

Acknowledgements

This work was supported by the University of Manchester Translational Imaging Unit and funded by the MIAS (Medical Images and Signals) IRC under EPSRC grant no. GR/N14248/01 and the UK Medical Research Council grant no. D2025/31, and by the Wellcome Trust grant no. ME003889. The software used in this study is freely available from www.tina-vision.net.

References

1. N. Kabani, G. Le Goualher, D. MacDonald et al. "Measurement of cortical thickness using an automated 3-D algorithm: A validation study." *NeuroImage* **13**, pp. 375–380, 2001.
2. V. A. Magnotta, N. C. Andreasen, S. K. Schultz et al. "Quantitative in vivo measurement of gyrification in the human brain: Change associated with aging." *Cerebral Cortex* **9**, pp. 151–160, 1999.
3. E. R. Sowell, P. M. Thompson, K. D. Tessner et al. "Mapping continued brain growth and grey matter density reduction in dorsal frontal cortex: Inverse relationships during postadolescent brain maturation." *J. Neurosci* **21**, pp. 8819–8829, 2001.
4. E. R. Sowell, B. S. Peterson, P. M. Thompson et al. "Mapping cortical change across the human life span." *Nature Neuroscience* **6**, pp. 309–315, 2003.
5. D. H. Salat, R. L. Buckner, A. Z. Snyder et al. "Thinning of the cerebral cortex in aging." *Cerebral Cortex* **14**, pp. 721–730, 2004.
6. A. W. Toga, P. M. Thompson & E. R. Sowell. "Mapping brain maturation." *Trends in Neurosciences* p. In Press, 2006.
7. P. M. Thompson, K. M. Hayashi, E. R. Sowell et al. "Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia." *NeuroImage* **23**, pp. S2–S18, 2004.
8. D. MacDonald, N. Kabani, D. Avis et al. "Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI." *NeuroImage* **12**, pp. 340–356, 2000.
9. S. E. Jones, B. R. Buchbinder & I. Aharon. "Three-dimensional mapping of cortical thickness using Laplace's equation." *Human Brain Mapping* **11**, pp. 12–32, 2000.
10. E. R. Sowell, P. M. Thompson, C. M. Leonard et al. "Longitudinal mapping of cortical thickness and brain growth in normal children." *J. Neurosci.* **24**, pp. 8223–8231, 2004.
11. E. Luders, K. L. Narr, P. M. Thompson et al. "Hemispheric asymmetries in cortical thickness." *Cerebral Cortex* p. In Press, 2006.
12. B. Fischl & A. M. Dale. "Measuring the thickness of the human cerebral cortex from magnetic resonance images." *PNAS* **97**, pp. 11050–11055, 2000.
13. J. S. Kim, V. Singh, J. K. Lee et al. "Automated 3-d extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification." *NeuroImage* **27**, pp. 210–221, 2005.
14. P. M. Thompson, A. D. Lee, R. A. Dutton et al. "Abnormal cortical complexity and thickness profiles mapped in Williams syndrome." *J. Neurosci.* **25**, pp. 4146–4158, 2005.
15. G. Lohmann, C. Preul & M. Hund-Georgiadis. "Morphology-based cortical thickness estimation." In C. J. Taylor & J. A. Noble (editors), *Information Processing in Medical Imaging, 18th International Conference, IPMI 2003, Ambleside, UK, July 20-25, 2003, Proceedings*, volume 2732 of *Lecture Notes in Computer Science*, pp. 89–100. Springer, 2003.
16. M. L. J. Scott & N. A. Thacker. "Regional cerebral cortical thickness measurement using MRI." In *Proc. MIUA*, pp. 7–10. 2005.
17. M. Pokrić, N. Thacker, M. L. J. Scott et al. "Multi-dimensional medical image segmentation with partial voluming." In *Proc. MIUA*, pp. 77–81. 2001.
18. J. Talairach & P. Tournoux. *Co-planner stereotaxic atlas of the human brain. 3-Dimensional proportional system: An approach to cerebral imaging*. Thieme Medical Publishers, New York, 1988.
19. J. F. Canny. "A computational approach to edge detection." *IEEE PAMI* **8(6)**, pp. 679–698, Nov 1986.
20. P. A. Bromiley, N. A. Thacker & A. Jackson. "Trends in brain volume change with normal ageing.", 2005. Proc. MIUA.
21. C. V. Economo. *The Cytoarchitectonics of the Human Cerebral Cortex*. Trasnlasted by S. Parker. Oxford University Press, Oxford, 1929.
22. D. Tosun, M. E. Rettmann, X. Han et al. "Cortical surface segmentation and mapping." *NeuroImage* **23**, pp. S108–S118, 2004.
23. R. Quester & R. Schroder. "The shrinkage of the human brain stem during formalin fixation and embedding in paraffin." *J. Neurosci. Methods* **75**, pp. 81–89, 1997.
24. N. C. Fox, S. Cousens, R. Scahill et al. "Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer's disease." *Archives of Neurology* **57**, pp. 339–344, 2000.