

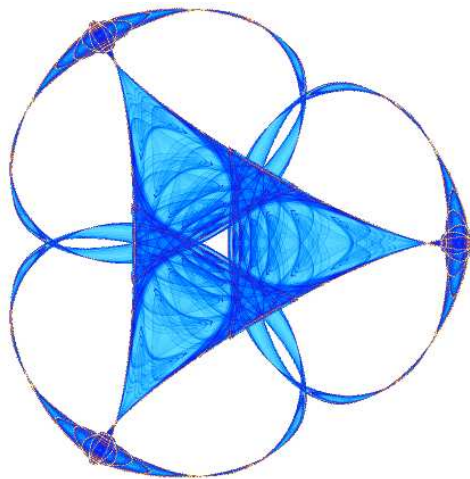
**SEGMENTATION-FREE MEASUREMENT OF CORTICAL THICKNESS
FROM MRI BY MINIMUM LINE INTEGRALS**

By

**Iman Aganj, Guillermo Sapiro
Neelroop Parikshak, Sarah K. Madsen
and
Paul M. Thompson**

IMA Preprint Series # 2196

(April 2008)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA
400 Lind Hall
207 Church Street S.E.
Minneapolis, Minnesota 55455-0436

Phone: 612/624-6066 Fax: 612/626-7370

URL: <http://www.ima.umn.edu>

Segmentation-Free Measurement of Cortical Thickness from MRI by Minimum Line Integrals

Iman Aganj, Guillermo Sapiro, *Senior Member, IEEE*, Neelroop Parikshak, Sarah K. Madsen, and Paul M. Thompson

Abstract—Estimating the thickness of the cerebral cortex is a key step in many brain imaging studies, revealing valuable information on development or disease progression. In this work, we present a framework for measuring the cortical thickness, based on minimizing line integrals over the probability map of the gray matter in the MRI volume. We first prepare a probability map which contains the probability of each voxel belonging to the gray matter. Then, the thickness is basically defined for each voxel as the minimum line integral of the probability map on line segments centered at the point of interest. In contrast to our approach, previous methods often perform a binary-valued segmentation of the gray matter before measuring the cortical thickness. Due to image noise and partial volume effects, such a hard classification ignores the underlying tissue class probabilities assigned to each voxel, discarding potentially useful information. We describe our proposed method and demonstrate its performance on both artificial volumes and real 3D brain MRI data from subjects with Alzheimer’s disease and healthy individuals.

Index Terms—Cortical thickness measurement, gray matter density, line integrals, magnetic resonance imaging, soft classification.

I. INTRODUCTION

MEASURING the cortical thickness has long been a topic of interest for neuroscientists. Cortical thickness changes in a characteristic pattern during childhood development and with the progression of neurodegenerative diseases such as Alzheimer’s, HIV/AIDS, and epilepsy [1], [2]. Recent studies examining changes in cortical thickness over time have revealed the trajectory of diseases in the living brain, and have been used to quantify treatment effects,

identifying regions where cortical thickness correlates with age, cognitive deterioration, genotype, or medication.

Various approaches have recently been proposed to automate this cortical thickness measurement from Magnetic Resonance Imaging (MRI) data, e.g., [3]–[10]. The limited spatial resolution of most MRI volumes (typically 1-2 mm) makes it difficult to measure cortical thickness accurately, which varies from 2 to 5 mm in different brain regions and is only a few voxels thick in the images. The neuroscience community has not yet agreed on a unique definition of cortical thickness and so far the various methods proposed measure slightly different quantities. What is common among them is that they virtually all perform a pre-segmentation of the white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), and most extract explicit models of the surfaces between them (i.e., the inner surface between WM and GM and outer surface between GM and CSF). They then use this hard segmentation as the input data for different tissue thickness measurement algorithms (Sec. II briefly reviews previous work). The disadvantage of this approach is that in the hard segmentation process, a considerable amount of information is discarded and never used in the measurement, not to mention the significant error in measured thickness that could be introduced by a few misclassified voxels (see Sec. IV.A for an example).

The approach we adopt here uses a soft pre-labeled volume as the input data. Due to the limited resolution of an MRI volume, many voxels contain partial amounts of two or more tissue types (see [11] and the references therein). Their intensity values give us information about the probability or proportion of those voxels belonging to any of the categories of WM, GM, or CSF. Rather than a pre-classified volume, we use one containing the probability that each voxel belongs to the GM.¹ These probability values have the same precision as the values in the original MRI volume, and therefore we do not discard any useful information.² We compute line integrals of the soft classified data, centered at each voxel and

Manuscript received ..., 2008. This work was supported in part by the National Institutes of Health (NIH), the National Science Foundation (NSF), the National Center for Research Resources (P41 RR14075), the National Institute for Biomedical Imaging and Bioengineering (R01 EB007813) and the Mental Illness and Neuroscience Discovery (MIND) Institute.

I. Aganj and G. Sapiro are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (email: {iman; guille}@umn.edu).

N. Parikshak, S. K. Madsen, and P. M. Thompson are with the Laboratory of Neuro Imaging, University of California-Los Angeles, School of Medicine, Los Angeles, CA 90095 USA (email: {neelroop.parikshak; sarah.madsen; thompson}@loni.ucla.edu).

¹ When considering partial volume effects, these “probabilities” represent the proportion of GM in the voxel.

² The only lost information is that we are not able to distinguish between WM and CSF, which as we will see is not a concern, since we are only interested in the two categories of GM and non-GM. However, we could also preserve that information if needed.

in all possible spatial directions, and then consider their minimum as the local cortical thickness at that voxel.

In Sec. II, we review previous work on cortical thickness measurement. Sec. III describes our proposed framework, and experimental results are presented in Sec. IV. Sec. V concludes with a review of the contributions, and finally the implementation is covered in detail in the Appendix.

An early conference version of this work was published in [12]. Here, we extend this work with more validation and comparisons. In addition, the implementation of the algorithm is provided in detail, and a new technique to find the GM skeleton is introduced.

II. PREVIOUS RELATED WORK

We now discuss some of the previously reported work for measuring the cortical thickness. Most methods require a pre-segmentation of the inner and outer surface, which results in a loss of available information and often inaccuracy in the input for the main thickness measurement algorithm.

Coupled surface methods [3], [13], define the cortical thickness as the Euclidean distance between corresponding point pairs on the inner and outer surfaces, often with parametric grids imposed. A displaced surface may result in an overestimation of the thickness (see Fig. 1(a)). *Closest point methods* such as [14] compute for each point on one of the two surfaces the closest point on the other surface and define the thickness as the distance between them. The main drawback with these methods is the absence of symmetry, as seen in Fig. 1(b). In another method introduced in [15], the *regional histogram of thickness* is estimated by measuring the length of the line segments connecting the inner and outer surfaces of the GM layer, normal to one of the surfaces. The median of the histogram is then chosen as the local cortical thickness. A detection of the WM-GM and GM-CSF boundaries is however necessary.

Laplace Methods [4], [7], [16] solve Laplace’s equation in the GM region with the boundary condition of constant (but different) potentials on each of the two surfaces. The cortical thickness is then defined on each point as the length of the integral curve of the gradient field passing through that point, as illustrated in Fig. 1(c). With this approach, the thickness is uniquely defined at every point. Nevertheless, a pre-segmentation of the two surfaces is required, reducing the accuracy of this technique.

Another category of methods defines thickness by making use of a central axis or skeleton [9], [17]. Thickness is typically estimated as the diameter of the largest enclosed sphere in the GM layer, which is (in some cases only initially) centered on a point on the central axis. As Fig. 1(d) demonstrates, a relatively sharp change in the thickness may result in a new branch and affect the topology of the skeleton.

The vast majority of the methods reported in the literature propagate segmentation errors to later steps, and segmentation is still in itself a challenging problem in brain imaging. Considering that the GM layer spans only a few voxels at the commonly used 1-2 mm resolutions, these errors can be

significant, and measuring tissue thickness avoiding this hard segmentation step may be very beneficial. This is the approach introduced here and described next.

III. METHODS

A. Definition

In its simplest form, we define the thickness of the GM at a given voxel as the minimum line integral of the probability map of the GM over all lines passing through that voxel. Formally:

$$T(\vec{x}) := \min_{l \in L_{\vec{x}}} \int_l P(\vec{r}) dl,$$

where $T(\vec{x})$ is the thickness of the GM at a point $\vec{x} \in \mathbb{R}^3$, $P(\vec{x})$ is the probability of the point \vec{x} belonging to the GM (estimation of this probability is described in Appendix), and $L_{\vec{x}}$ is the set of all lines in three-dimensional space passing through the point \vec{x} . In practice, however, $L_{\vec{x}}$ is comprised of all equal-length *line segments* centered at \vec{x} , which are sufficiently longer than the expected maximum thickness in the volume. Choosing longer line segments does not greatly affect the integral values since $P(\vec{r})$ decreases significantly on the non-GM regions.³ Fig. 2(a) shows an example of this construction, for a 2D binary probability map, where the probability of belonging to the GM is 1 inside the shape and 0 outside. When computing thickness at the specified point, the line segment marked with oval arrows is selected as the one giving the smallest line integral. The corresponding integral value, which in this case is the length of its overlap with the GM (in bold), is the thickness of the GM at that point. A more realistic situation is shown in Fig. 2(b), where the probability map varies between zero and one. A blurred border, which results from the limited resolution of the MRI, includes voxels that partially contain GM. Due to the pre-segmentation, this type of partial volume information is not considered in most prior work in this area.⁴

Our method is based on an intuitive way of measuring the thickness of an object. A simple way to measure the local thickness of an object would be to put two fingers on both edges of the object, and move the finger tips locally (equivalent to varying the angle of the segment connecting them to each other), until the distance between them is minimized (Fig. 3). For example, among Figs. 3(a), 3(b), and 3(c), we would naturally choose Fig. 3(b) as the one that depicts the most accurate local thickness. This distance could then be considered as the local thickness of the object. Thus, we are dealing with a constrained optimization problem: minimizing a distance in a specific region. In our approach,

³ When computing $T(\vec{x})$ at a non-GM voxel, the value will be zero due to the stopping criteria explained below.

⁴ The proposed framework is independent of how the probability density functions are actually computed. The key is to use the probabilities instead of the hard thresholds. The computation of these probabilities for the examples in this paper is detailed in the Appendix.

however, this region is identified precisely by the point where we want to define the thickness. Therefore, the constraint is that the point must be on the line segment connecting the two finger tips. In other words, we consider only the line segments passing through the point where we intend to find the thickness. The minimized distance – or the length of the line segment – is in this case the integral of the probability map on the line containing the segment (Fig. 3(e)).

B. Algorithm

The algorithm basically computes every line integral centered at each point of the volume starting from the point of interest and proceeding in each of the two opposite directions separately (see Appendix for our discretization method). Once all the line integrals at a point are calculated (meaning in all possible directions), the minimum of them is considered to be the thickness at that point. However, to reduce the effect of noise, an alternative would be for instance to consider the average of some of the smallest integrals.

In practice, a problem may arise, typically in narrow sulci where the outer surface of the folded GM layer has two nearby sides (Fig. 4(a)). While computing the thickness on one part of the layer, the GM of the other part may be partially included in some of the line integrals; this will lead to the thickness being overestimated (Fig. 4(b)). To avoid this error, we include two stopping criteria which prevent a line integral from further advancing when it is believed that no more summation is necessary or that we are mistakenly considering a different region of the GM layer. The line integral stops proceeding if the probability map:

1. Has been below a specific threshold for a certain number of consecutive voxels, or
2. Has been decreasing at least for a certain number of successive voxels and then increasing for an additional number of voxels.

We use the first criterion, since if the probability has been low for a while, we are most likely not in the GM region anymore, and by further summing we would just increase the error. An additional advantage of using this stopping criterion is that summation will be stopped quickly after starting to measure the thickness based on voxels that are not in the GM region. The algorithm will ignore those points and will return near-zero values for the GM thickness at those locations.

The second condition happens when two parts of the GM layer are so close to each other that the probability on the gap between them is not small enough for the first stopping condition to become true, therefore the algorithm stops summing after identifying a valley on the probability map. As we see in the Appendix, the algorithm can be implemented such that gaps as narrow as one voxel are detected by the above stopping criteria.⁵

⁵ In case the gap is narrower, both GM layers basically touch each other and the algorithm may overestimate the thickness. Restricting the thickness map to be continuous, which is a part of the future work, may help us overcome this problem.

IV. RESULTS AND DISCUSSION

A. Artificial Data

To illustrate and validate our approach, we first show results using artificial input data. Fig. 5(a) shows the isosurfaces of an artificially created probability map of a parabolic-shaped layer of GM with varying thickness in a volume of $50 \times 50 \times 50$ voxels. The two isosurfaces represent the inner and outer gray matter surfaces. Depicted as small circles, a number of sample points have been selected, where the computed thicknesses are illustrated as line segments. The direction of each line segment is the optimal direction that gives the minimum line integral of the probability map. The thickness is indicated in the figure by the length of the line segments.

Next, to show the negative consequences of hard segmentation when noise and partial volume effects are present, we reduced the resolution of the volume five times by taking the mean value of every $5 \times 5 \times 5$ sub-volume; we also added zero-mean Gaussian noise with standard deviation of 0.2 (Fig. 5(c)), and ran the measurement algorithm on it. In addition, we performed hard segmentation on the low-resolution, noisy volume, by substituting the probability values less than 0.5 with 0 and other values with 1 (Fig. 5(d)), and re-ran the measurement algorithm. Using the results of the original high-resolution case as ground truth, the experiments on the low resolution and noisy volumes showed an average error in the estimated thickness of 1.9 voxels in the segmentation-free case and 2.2 voxels when hard segmentation was performed as a pre-step for reporting this measurement.

B. Real MRI Data

We tested the proposed technique on 44 T1-weighted brain MRI scans, acquired using a standard sagittal 3D MP-RAGE sequence (TR: 2400 ms, minimum TE, inversion time (TI) 1000 ms, flip angle: 8° , 24 cm field of view), with a reconstructed voxel size of $0.9375 \times 0.9375 \times 1.2 \text{ mm}^3$. To adjust for scanner – and session – specific calibration errors, standard corrections were made for gradient nonlinearity, phantom-based scaling, and adjustment of intensity inhomogeneity [18]. The actual tissue probabilities $P(\vec{r})$ needed in our method are computed following the procedure explained in the Appendix.

A 2D slice from an MRI volume is shown in Fig. 6(a) along with its computed thickness map in Fig. 6(b). Since we do not extract the GM, the results also contain thickness values for other parts of the head such as the scalp, which may be ignored. Fig. 6(c) illustrates a 3D surface-based mapping of the cortical thickness visualized by the mrGray software, using the steps in [19].

Our dataset includes pairs of scans over a one-year interval from 22 subjects (total of 44 scans), of whom 9 had been diagnosed with Alzheimer’s disease at their first scan, and 13 were age-matched normal subjects. These subjects were included in one of our prior morphometric studies, where the

scanning protocol is detailed [20]. Each subject was scanned at 1.5 Tesla with a 3D T1-weighted acquisition, with the following parameters: repetition time (TR), 2400 ms; minimum full TE; inversion time (TI), 1000 ms; flip angle, 8°, 24 cm field of view, yielding a reconstructed voxel size of $0.9375 \times 0.9375 \times 1.2 \text{ mm}^3$. The images were calibrated with phantom-based geometric corrections to ensure stable spatial calibration over time [21].

For comparison,⁶ we also analyzed our data using the FreeSurfer thickness computation technique, [3]. The results of the change in the mean thickness over a year are demonstrated for the individual cases in Fig. 7, while the corresponding detailed statistical data can be seen in Table I. From the computational point of view, our approach is 1-2 orders of magnitude faster than the one implemented in FreeSurfer, and it is highly parallelizable.

Both techniques are able to detect a systematic change in cortical thickness over time, both in the AD group and in controls. Our method found thickness declined in AD ($t_{paired} = -1.200$; $p=0.001$), and in controls ($t_{paired} = -0.274$; $p=0.009$); FreeSurfer found thickness declined in AD ($t_{paired} = -1.495$; $p=0.001$), and increased in controls ($t_{paired} = +0.175$; $p=0.001$).⁷

First of all, the thickness measures are extremely highly correlated between baseline and follow-up scan, for both measurement methods and both subject groups. Our technique gives measures that correlated highly over the 1 year time-interval in AD ($p=0.001$) and in controls ($p=0.009$), and so did FreeSurfer ($p=0.001$ both for AD and for normal subjects).

Second, although changes over time are more difficult to measure than absolute values of thickness, our technique correlates highly with FreeSurfer in the changes it measures (Pearson's $r=0.572$; $p=0.0028$). This agreement between methods was also found in the group of healthy controls, where changes are minimal ($r=0.583$; $p=0.0011$), and in the group of AD patients ($r=0.558$; $p=0.0047$).

The changes over time that were detected by FreeSurfer tended to be a little lower in general than those detected by our method, and satisfied the following least-squares regression model: $\Delta\text{thickness}_{FS} \sim [0.45 * \Delta\text{thickness}_{OurMethod}] - 0.04$.⁸ Even so, a Student's t test designed to compare thickness values across the two methods did not recover any systematic biases between methods ($t_{paired}=0.557$, $p>0.05$). Our method gives a slightly higher SD in the thickness measures at each time point, and this, together with the fact that these errors may be uncorrelated over time, may also lead to a slightly higher estimation of change. Even so, if the sources of

errors/deviations are uncorrelated across subjects, there should be no bias incurred by using one method versus the other.

A number of observations can be made from these experiments and the comparison with FreeSurfer. While the actual exact value of cortical thickness decline is not known for the AD subjects, for biological reasons a net increase is not expected for either of the two populations. This is partly because cortical thinning is a natural process that occurs due to neuronal shrinkage in normal aging, and is accelerated in AD due to cell death and neuronal loss in the cortex. Thereby, any apparent increase points to the difficulty and measurement errors in cortical thickness methods. Both our method and FreeSurfer report a few subjects that show increases, while on average our proposed method shows decline for both populations (much more significant for the AD subjects, as expected). Even so, FreeSurfer shows a net increase for the control subjects, not very large but still statistically significant. A number of reasons might explain this, including registration issues and segmentation errors, which are minimized in our proposed technique. Due to the existence of such intrinsic difficulties in measuring cortical thickness, the simultaneous running of different algorithms, with an evaluation of any inconsistencies or consensus, is an important alternative.

C. Correlation with Clinical Data

In addition to classification accuracy, it is desirable that any measure of cortical thickness can be shown to be associated with clinical measures of deteriorating brain function. This is because image-derived measures often serve as a proxy for measures of disease burden that are based on repeated cognitive tests, repeated pathological tests (lumbar puncture). When we modeled factors that affected rates of thinning in AD, our method found a significant sex difference in the rate of thinning ($r=0.428$; $p<0.037$), whereas FreeSurfer detected this sex difference at a trend level ($r=0.389$; $p<0.060$), with women experiencing a more rapid rate of loss. Overall, the correlations with clinical measures in AD, such as the Mini-Mental State Exam scores at baseline and 6 month follow-up, and the changes in those scores, were around $r=0.37-0.5$ for FreeSurfer ($p=0.013-0.073$), and slightly lower for our method ($r=0.24-0.30$; $p>0.1$). Even so, this may be due to the slightly higher standard deviation for the changes reported by our method. Conversely, our method detected associations between the cortical thinning rate and Geriatric Depression Scores ($r=0.433$; $p=0.0215$), but FreeSurfer did not detect such an association ($r=0.106$; $p>0.1$). Clearly, a head-to-head comparison on a larger sample would be useful, including, for example, assessments of the sample sizes needed by both approaches to detect a 25% slowing of the disease with 80-90% power.

V. CONCLUSION

We presented a new definition of cortical thickness along with an algorithm for computing it. We were motivated by the importance of measuring the thickness of the cerebral cortex for quantifying the progression of various neurodegenerative brain diseases. Our method calculates the thickness at each

⁶ See Appendix D for a detailed explanation about what voxels are used to perform this comparison (recall that our algorithm produces thickness measurements for all voxels).

⁷ This increase, which is anatomically not expected, is statistically significant. No increase is found with our proposed method.

⁸ We should note that while this is the best fitting straight line, in a least squares sense, the actual fitting is far from ideal, indicating that the relation between the measurements produced by both methods is not linear. Note of course that intrinsically our definition of tissue thickness is fundamentally different than that of FreeSurfer, and should be considered as an alternative.

voxel, by computing all line integrals of the probability map of the gray matter passing through that voxel, and choosing the minimum integral value. Two stopping criteria are taken into consideration to address issues created by narrow sulci. Unlike most prior work, we take into account the probability of each voxel belonging to the gray matter layer and do not carry out a hard segmentation prior to measuring the thickness. The proposed algorithm is significantly faster than popular ones such as the one implemented in FreeSurfer. We have validated the technique with artificial data and presented reasonable results for longitudinal MRI scans of Alzheimer’s disease and normal subjects.

We are currently investigating the incorporation of smoothness constraints, where the minimum line integral at a given voxel is influenced by those of neighborhood voxels. We are also applying the framework introduced here to large population studies.

APPENDIX

We now explain in detail how we implemented our proposed algorithm. The input to the algorithm is the raw MRI dataset which is in the form of a three-dimensional matrix, and the result is the thickness map volume with the same size and spatial sampling as the input.

A. Computing the Probability Map

We consider the probability of each voxel belonging to the GM as a Gaussian distribution on the intensity value of the voxel in the MRI volume. The mean of the Gaussian is the mean value of manually-selected sample voxels in the GM, while the standard deviation is the difference between the manually estimated mean values of GM and WM. We could use more sophisticated soft classification algorithms, such as Partial-Volume Bayesian algorithm (PVB) [22], Probabilistic Partial Volume Classifier (PPVC) [23], and Mixel classifier [24], to further improve the results. Before segmentation, several pre-processing steps were applied to ensure the accurate calibration of the scans over time and across subjects. Specifically: (1) a procedure termed *GradWarp* was applied for correction of geometric distortion due to gradient non-linearity [25], (2) a “B1-correction” was applied, to adjust for image intensity non-uniformity using B1 calibration scans [21], (3) “N3” bias field correction, for reducing intensity inhomogeneity caused by non-uniformities in the radio frequency (RF) receiver coils [26], and (4) geometrical scaling, according to a phantom scan acquired for each subject [21], to adjust for scanner – and session – specific calibration errors. In addition to the original uncorrected image files, images with all of these corrections already applied (*GradWarp*, B1, phantom scaling, and N3) are available to the general scientific community (www.loni.ucla.edu/ADNI).

B. Preparing the Masks

In order to compute the line integrals, we made separate cubic mask volumes for all the line segments which are in different quantized directions in three-dimensional space. The length of each dimension of the mask volumes was chosen to be slightly longer than the actual length of the line segment. However, since binary masks are highly inaccurate to apply in numerical integration, we considered continuous mask volumes. We first built binary masks four times bigger in each dimension, and instead of line segments, we considered cylinders four times longer than the line segments with diameter four. Next, we applied a low pass filter to it and downsampled the results to achieve the desired non-binary masks. They were afterwards normalized so that the values in each mask added to the length of the line segment. A two-dimensional example is depicted in Fig. 8.

The directions were chosen by quantizing the unit hemisphere uniformly, as the complete sphere would have been redundant. In the standard spherical coordinates, the unit hemisphere was sampled every 10° in the latitudinal direction, whereas the longitudinal sampling rate varied as follows:

$$\begin{aligned} \theta &\in [0^\circ, 90^\circ] && \text{every } 10^\circ \\ \phi &\in [0^\circ, 360^\circ] && \text{every } 10^\circ / \cos \theta \end{aligned}$$

This particular quantization was chosen in order to achieve a relatively constant surface sampling rate, since the solid angle would be

$$\Delta\Omega = \cos \theta \Delta\theta\Delta\phi = \left[\cos \theta \cdot 10^\circ \cdot \frac{10^\circ}{\cos \theta} \right]_{\text{in steradians}} = \text{const.}$$

In addition, each mask contained only half of its corresponding line segment (extending from the origin to one point on the hemisphere), given that the two halves were identical and also that the line integrals on them would later need to be taken separately.

Next, the mask volumes were stored in a sparse format which is a list of non-zero elements with their coordinates. To facilitate the integration, elements were sorted into ascending order with respect to their Euclidean distance from the origin. The masks were then stored for further use in the numerical integration step, for various input data.

C. Numerical Integration

We now briefly explain how the line integrals of a probability map are numerically calculated. To compute the thickness measure at each given voxel, we integrate the probability map on line segments centered at that voxel in different directions using the pre-computed mask volumes, and then choose the smallest integral value as the local thickness. Since our stopping criteria make the entire operation nonlinear, we are not able to use convolution methods – which are computationally faster if done in frequency domain, although demanding more memory – to compute the line integrals.

Each line integral is computed in two similar phases. Starting from the main voxel for which we intend to compute the thickness, in each phase we advance in one of the two opposite directions, using the mask volume in the first phase, and its symmetry about the origin (with negated coordinates) in the second phase. In each phase, we start adding the probability values on the region around the main voxel using as weights the values in the sparse form of the corresponding mask. Since the non-zero elements of the mask are sorted into ascending order with respect to their Euclidean distance to the origin, the values of the closer voxels to the main voxel are added first.

The line integral does not need to be entirely computed, since we have a record of the minimum integral value that has been computed so far for a voxel, and we do not need to exceed that. Therefore, the summation continues until one of the following occurs:

1. All the non-zero elements of the mask are used.
2. The sum reaches or exceeds the minimum integral value that has so far been computed for that voxel.
3. One of the stopping criteria becomes true (see Sec. III.B).

For the first stopping condition (explained in Sec. III.B), a counter is set for the number of successive voxels (with the original sampling size) encountered with probabilities less than the threshold. In this work we used the threshold of 0.3, and summing stopped as the counter reached 10. For the second condition, two states are necessary. In the initial state, a variable counts the number of consecutive decreases in the probability. When it reaches a specific number, the state changes. In the second state, a new variable counts the number of increases in the probability and the summation stops when it reaches a predefined number. In this work 10 was used for both mentioned numbers. Since the masks are made in an interpolative way, they each have many more non-zero voxels than the real length (in voxels) of their corresponding line segments. Therefore, the number of counted iterations is usually about ten times larger than the actual proceeded length, which results in higher precision in counting. For example, a one-voxel thick valley can be detected after about five consecutive decreases and five consecutive increases of the probability.

The process of measuring and comparing the thickness for a pair of scans took about 2 to 3 hours in our technique, in contrast to 2 to 3 days running time of FreeSurfer. In both cases, a Linux machine with a 2.33GHz Intel CPU was used. One can also parallelize the above procedure, since it is done independently for each voxel in the volume. For instance, we divided the volume into eight sub-volumes and ran eight parallel jobs in different processors to obtain the thickness map in considerably less time.

D. Finding the Skeleton

Since the output of the proposed algorithm is the set of thickness values for all the voxels in the volume, we need to find the particular voxels that lie inside the GM layer to compute statistical results for group comparisons, and for

comparing the results with FreeSurfer.⁹ A simple GM mask may not necessarily represent the most appropriate set of voxels to examine. Not only might the thickness be underestimated at the voxels close to the inner/outer surfaces, but the values would also be weighted depending on how thick each segment is, i.e., more voxels with the same thickness value are counted on the thicker parts of the GM. Thus, it is best to make a *skeleton* of the GM, which is a mask with constant thickness positioned in the middle of the layer, far from the inner/outer surfaces.

In order to find the skeleton without explicit segmentation, we modified the algorithm such that for each voxel v , in addition to the total thickness $t(v)$, it reports the two different thickness values $t_1(v)$ and $t_2(v)$, the lengths of the two sub-segments on each side of the voxel on the optimal line segment ($t(v) = t_1(v) + t_2(v)$, see Fig. 9). Next, we considered the skeleton to be the set of the voxels with the following properties:

$$\begin{aligned} |t_1(v) - t_2(v)| &\leq 0.5, \\ P(v) &> 0.4, \end{aligned}$$

where $P(v)$ is the probability of v belonging to the GM. The first condition guarantees that the skeleton is always at most one voxel thick, and the two conditions together guarantee that the skeleton remains in the middle of the GM layer.

In addition, to better remove the skull and non-GM parts of the brain, we applied a rough GM mask obtained from FreeSurfer.

REFERENCES

- [1] P. M. Thompson *et al.*, "Abnormal cortical complexity and thickness profiles mapped in Williams syndrome," *J. Neuroscience*, vol. 25, no. 16, pp. 4146–4158, Apr. 2005.
- [2] P. M. Thompson *et al.*, "Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia," *NeuroImage*, vol. 23, pp. 2–18, Sep. 2004.
- [3] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," *Proc. Nat. Acad. Sci.*, vol. 97, no. 20, pp. 11 050–11 055, 2000.
- [4] S. E. Jones, B. R. Buchbinder, and I. Aharon, "Three-dimensional mapping of the cortical thickness using Laplace's equation," *Hum. Brain Mapping*, vol. 11, pp. 12–32, 2000.
- [5] N. Kabani, G. Le Goualher, D. MacDonald, and A. C. Evans, "Measurement of cortical thickness using an automated 3-D algorithm: a validation study," *NeuroImage*, vol. 13, no. 2, pp. 375–380, Feb. 2001.
- [6] G. Lohmann, C. Preul, and M. Hund-Georgiadis, "Morphology-based cortical thickness estimation," *Inf Process Med Imaging*, vol. 2732, pp. 89–100, 2003.
- [7] A. J. Yezzi and J. L. Prince, "An Eulerian PDE approach for computing tissue thickness," *IEEE Trans. Med. Imag.*, vol. 22, pp. 1332–1339, Oct. 2003.
- [8] J. P. Lerch and A. C. Evans, "Cortical thickness analysis examined through power analysis and a population simulation," *NeuroImage*, vol. 24, pp. 163–173, 2005.
- [9] N. Thorstensen, M. Hofer, G. Sapiro, H. Pottmann, "Measuring cortical thickness from volumetric MRI data," unpublished, 2006.
- [10] K. Young, N. Schuff, "Measuring structural complexity in brain images," *NeuroImage*, to be published.
- [11] D. L. Pham and P. L. Bazin, "Simultaneous boundary and partial volume estimation in medical images," in *Proc. 7th Intl. Conf. Medical Image*

⁹ Note once again that such segmentation is not an intrinsic component of the proposed framework, it is just one way of reporting the results.

- Computing and Computer-Assisted Intervention*, Saint-Malo, 2004, pp. 119–126.
- [12] I. Aganj, G. Sapiro, N. Parikshak, S. K. Madsen, and P. Thompson, "Segmentation-free measurement of cortical thickness from MRI," in *Proc. Fifth IEEE Int. Symp. on Biomed. Imag.*, Paris, France, 2008.
- [13] D. MacDonald, N. Kabani, D. Avis, and A. C. Evans, "Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI," *NeuroImage*, vol. 12, pp. 340–356, 2000.
- [14] M. I. Miller, A. B. Massie, J. T. Ratnanather, K. N. Botteron, and J. G. Csernansky, "Bayesian construction of geometrically based cortical thickness metrics," *NeuroImage*, vol. 12, pp. 676–687, 2000.
- [15] M. L. J. Scott and N. A. Thacker, "Cerebral cortical thickness measurements," Imaging Science and Biomedical Engineering Division, University of Manchester, Manchester, England, TINA Memo 2004-007, 2004.
- [16] H. Haidar, J. S. Soul, "Measurement of cortical thickness in 3D brain MRI data: Validation of the Laplacian method," *NeuroImage*, vol. 16, pp. 146–153, 2006.
- [17] S. Pizer, D. Eberly, D. Fritsch, and B. Morse, "Zoom-invariant vision of figural shape: The mathematics of cores," *Comput. Vision Image Understanding*, vol. 69, no. 1, pp. 55-71, 1998.
- [18] A. D. Leow *et al.*, "Longitudinal stability of MRI for mapping brain change using tensor-based morphometry," *NeuroImage*, vol. 31, pp. 627-640, 2006.
- [19] P.C. Teo, G. Sapiro and B.A. Wandell, "Creating connected representations of cortical gray matter for functional MRI Visualization," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 852-863, 1997.
- [20] X. Hua *et al.*, "3D characterization of brain atrophy in Alzheimer's disease and mild cognitive impairment using tensor-based morphometry," *NeuroImage*, published online, Feb. 2008.
- [21] C. Jack *et al.*, "The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI Methods," *Journal of MRI*, vol. 27, pp. 685-940, Apr. 2008.
- [22] D. H. Laidlaw, K. W. Fleischer, and A. H. Barr, "Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms," *IEEE Trans. Med. Imag.*, vol. 17, pp. 74-86, Feb. 1998.
- [23] H. S. Choi, D. R. Haynor, and Y. M. Kim, "Multivariate tissue classification of MRI images for 3-D volume reconstruction—A statistical approach," in *Proc. SPIE Medical Imaging III: Image Processing*, 1989, vol. 1092, pp. 183-193.
- [24] H. S. Choi, D. R. Haynor, and Y. M. Kim, "Partial volume tissue classification of multichannel magnetic resonance images—A mixel model," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 395-407, 1991.
- [25] J. Jovicich *et al.*, "Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data," *Neuroimage*, vol. 30, pp. 436-443, 2006.
- [26] J. G. Sled, A. P. Zijdenbos, and A. C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, pp. 87-97, Feb. 1998.

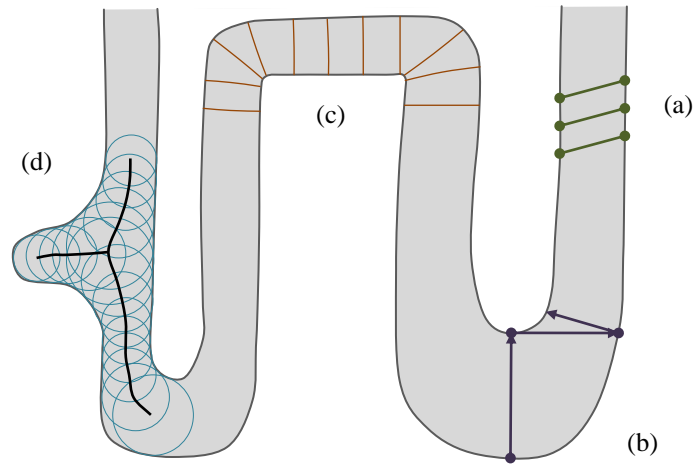
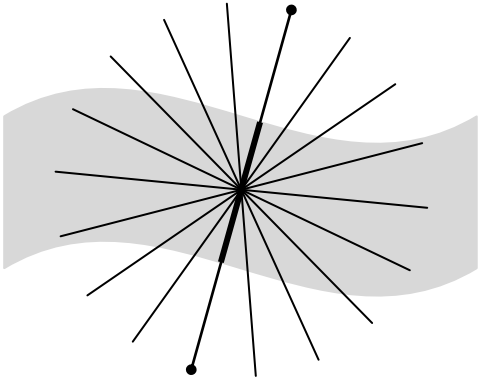
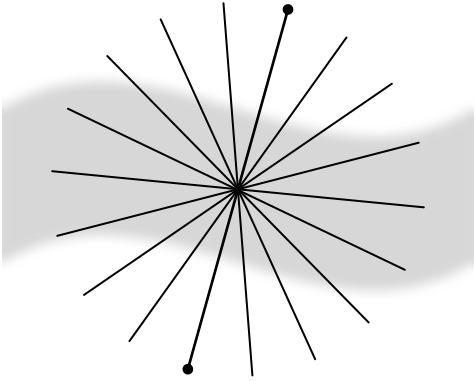


Fig. 1. Common ways of measuring cortical thickness. (a) Coupled surface methods. (b) Closest point methods. (c) Laplace ('heat-flow') methods. (d) Largest enclosed sphere methods.



(a)



(b)

Fig. 2. Computing line integrals passing through a point, and choosing the minimum integral value as the thickness. (a) Binary probability map. (b) Continuous probability map.

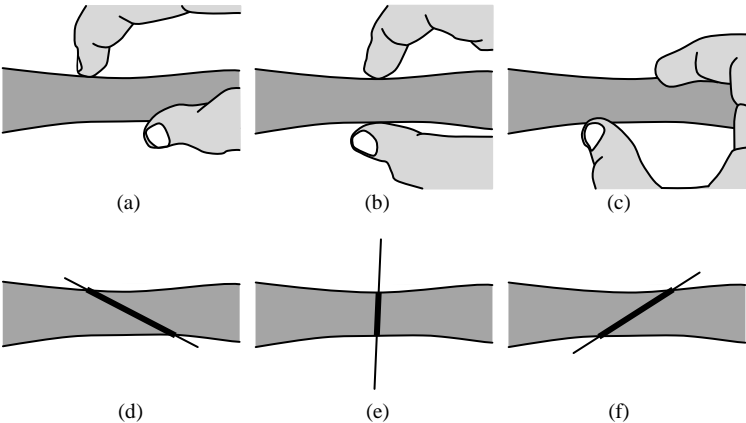
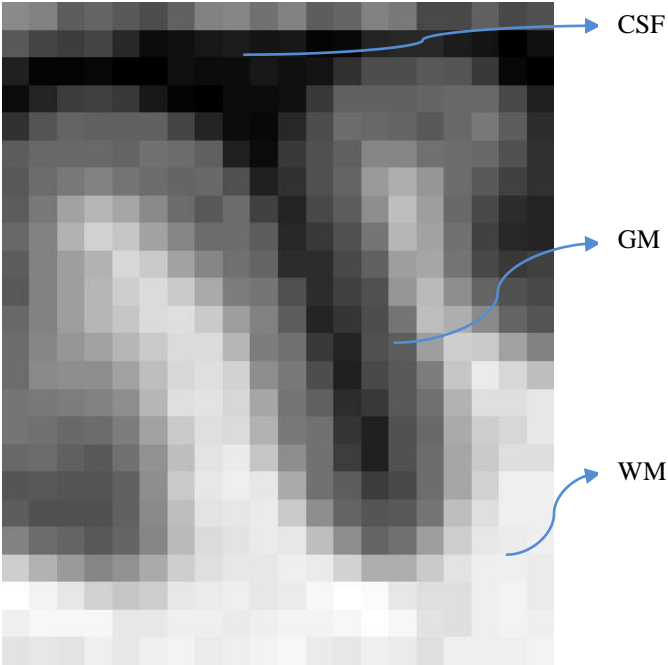
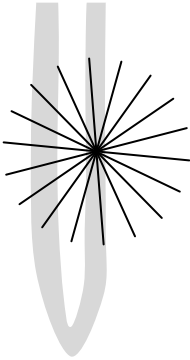


Fig. 3. (a), (b), (c) An intuitive way of measuring the thickness of an object. (d), (e), (f) Our algorithm produces results similar to this intuitive approach.

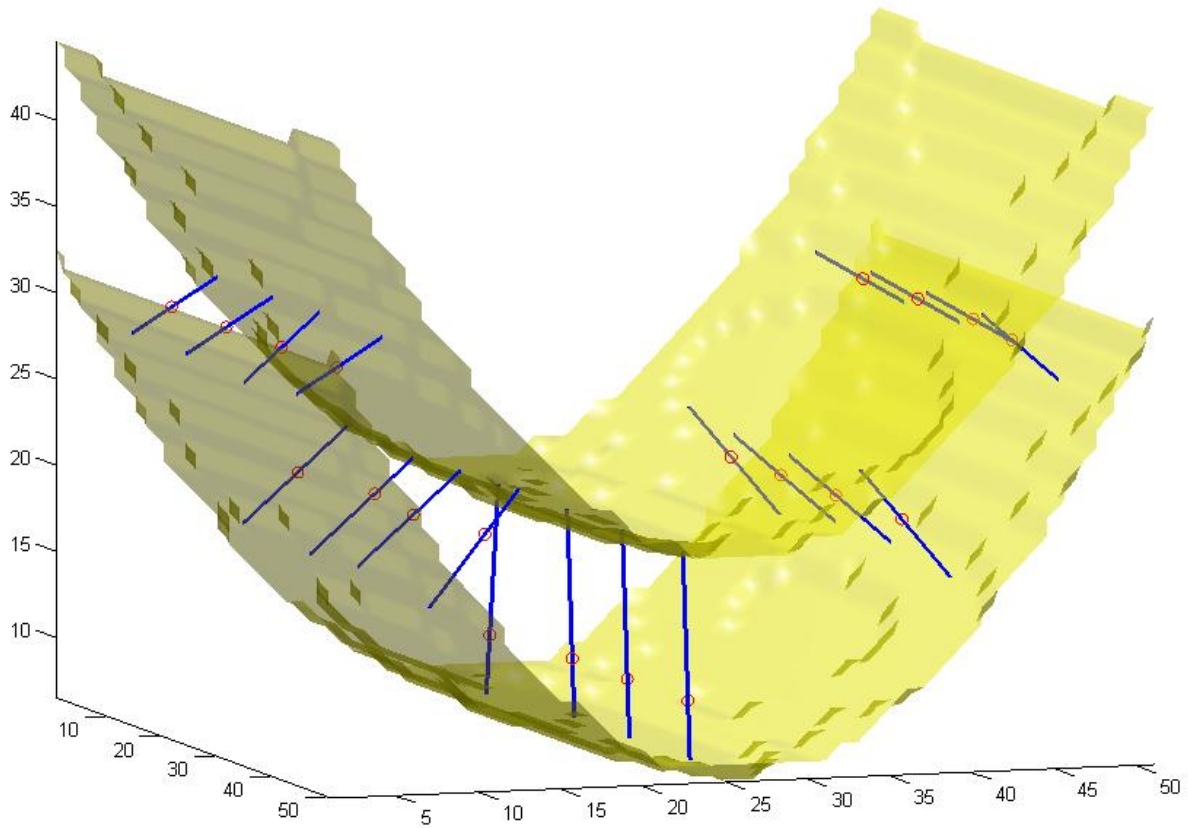


(a)

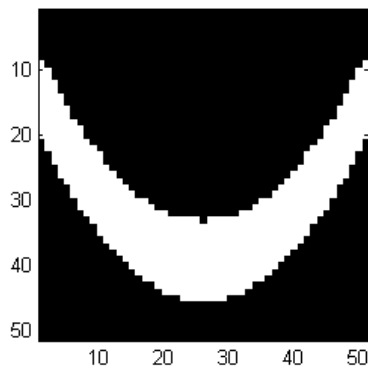


(b)

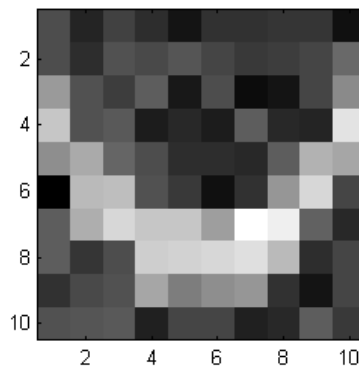
Fig. 4. (a) A sulcus in which two sides of the gray matter layer are close to each other. (b) How the algorithm might overestimate the thickness if no stopping criteria were used.



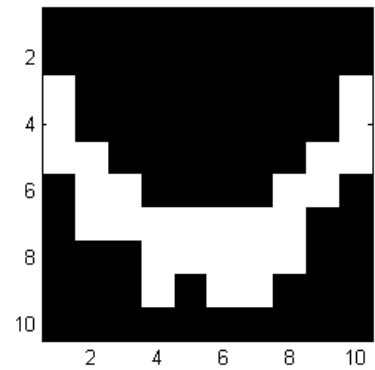
(a)



(b)



(c)



(d)

Fig. 5. Results on an artificial probability map. (a) Inner and outer surfaces of a parabolic-shaped layer of “GM” are depicted. Line segments are chosen by the algorithm such that they give the smallest integrals (of the probability map) among all line segments passing through every selected test point, here shown as small circles. (b) A 2D slice of the volume. (c) The same slice in a five-times-lower-resolution volume with additive Gaussian noise. (d) Binary classification of the low-resolution volume.

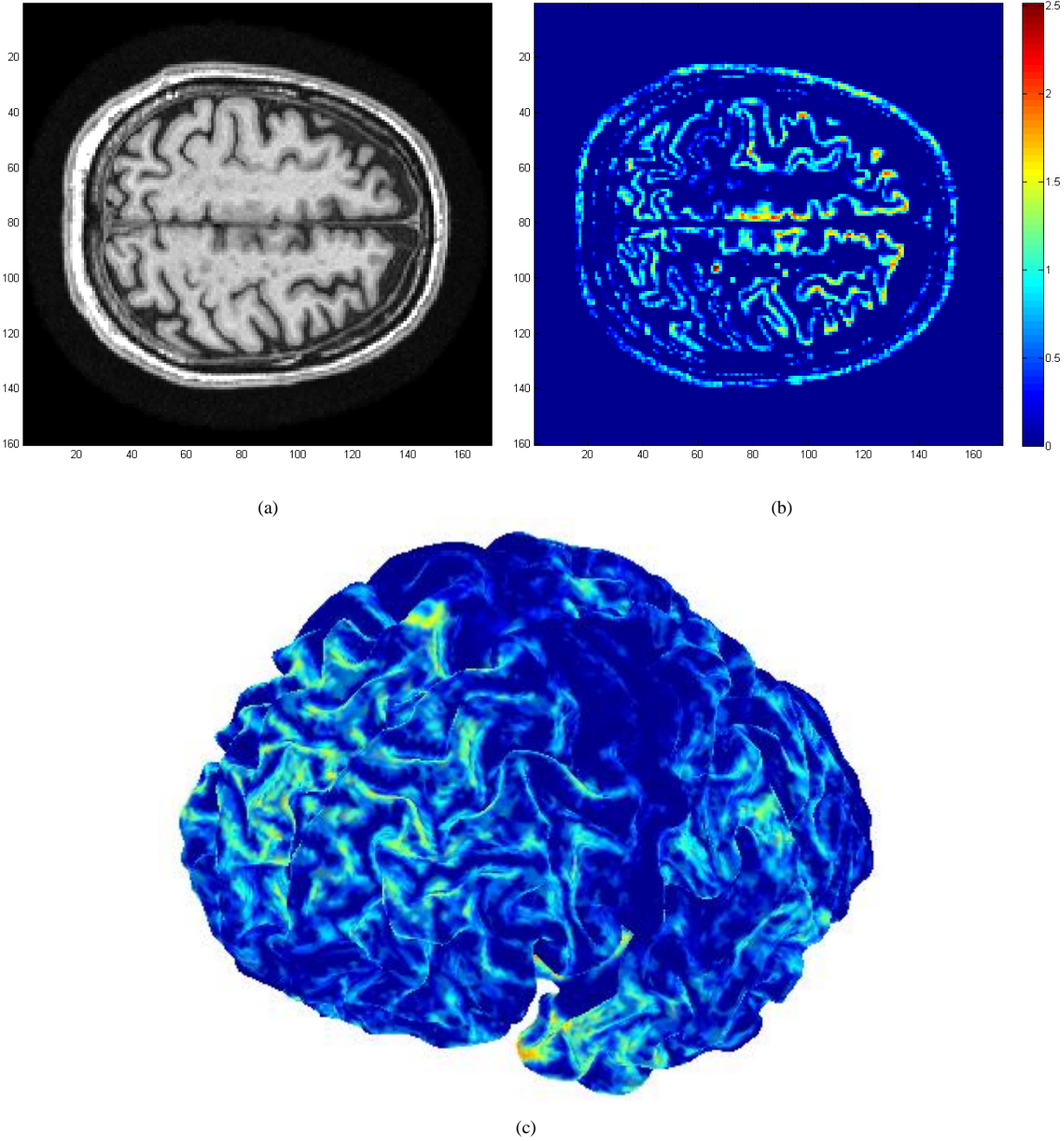
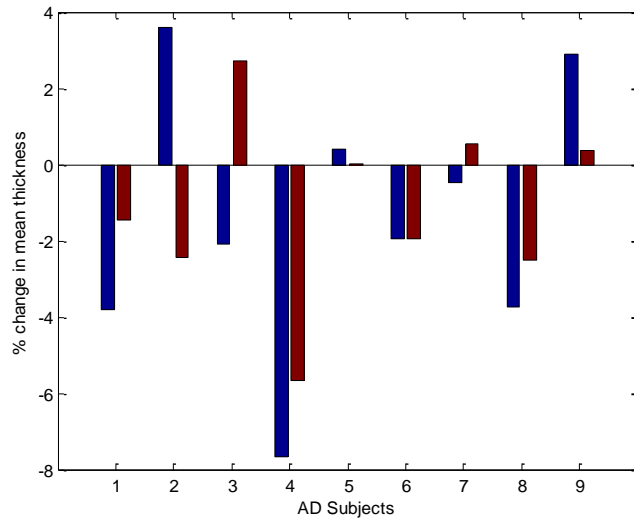
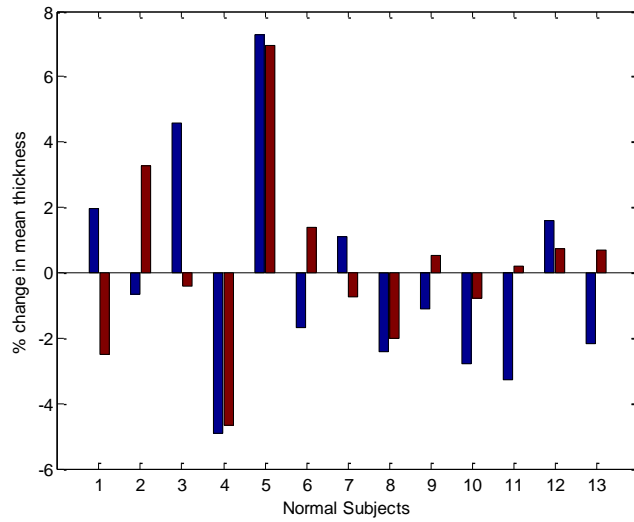


Fig. 6. Experimental results on MRI data. All computations were done in 3D. (a) A slice of the original volume. (b) The thickness map of the same slice (blue thinner, red thicker). (c) 3D mapping of the cortical thickness.



(a)



(b)

Fig. 7. Relative change in the mean cortical thickness over a one-year interval, comparing our results (left blue bars) with the results from FreeSurfer (right red bars). (a) Subjects diagnosed with Alzheimer's disease (AD). (b) Normal elderly subjects.

TABLE I
STATISTICAL RESULTS ON REAL MRI DATA

Quantity	Proposed Algorithm	FreeSurfer
Mean change for all subjects	-0.69 %	-0.35 %
Mean change for AD subjects	-1.42 %	-1.15 %
Mean change for normal subjects	-0.19 %	+0.20 %
SD of change for all subjects	3.42	2.68
SD of change for AD subjects	3.52	2.40
SD of change for normal subjects	3.39	2.82
Group separation	0.36	0.50

The relative change in the mean thickness as measured over a one year interval.
Group separation = (Mean for normal - Mean for AD) / (SD for all)

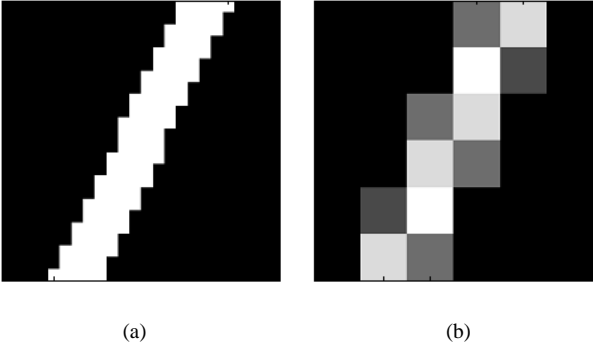


Fig. 8. Illustration of how line integral masks are generated. (a) A high-resolution binary mask, which is four times larger in each direction, is first generated. (b) The real-size non-binary mask is produced by downsampling the binary mask.

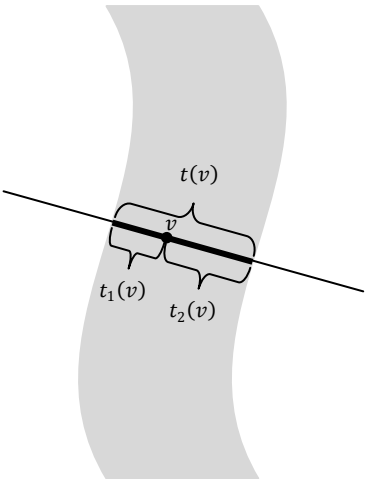


Fig. 9. The lengths, $t_1(v)$ and $t_2(v)$, of the two sub-segments are used in finding the gray matter skeleton for computing the cortical thickness variation.