Published in final edited form as:

Inf Process Med Imaging. 2017 June ; 10265: 28–40. doi:10.1007/978-3-319-59050-9_3.

Kernel Methods for Riemannian Analysis of Robust Descriptors of the Cerebral Cortex

Suyash P. Awate¹, Richard M. Leahy², and Anand A. Joshi²

¹Computer Science and Engineering Department, Indian Institute of Technology (IIT) Bombay, Mumbai, India

²Signal and Image Processing Institute (SIPI), University of Southern California, Los Angeles, USA

Abstract

Typical cerebral cortical analyses rely on spatial normalization and are sensitive to misregistration arising from partial homologies between subject brains and local optima in nonlinear registration. In contrast, we use a descriptor of the 3D cortical sheet (jointly modeling folding and thickness) that is robust to misregistration. Our histogram-based descriptor lies on a *Riemannian manifold*. We propose new *regularized nonlinear methods* for (i) detecting group differences, using a *Mercer kernel* with an implicit lifting map to a reproducing kernel Hilbert space, and (ii) regression against clinical variables, using *kernel density estimation*. For both methods, we employ kernels that exploit the Riemannian structure. Results on simulated and clinical data shows the improved accuracy and stability of our approach in cortical-sheet analysis.

Keywords

Cortex; Folding; Thickness; Robustness; Riemannian space; Reproducing kernel hilbert space; Hypothesis tests; Kernel regression

1 Introduction

Cerebral cortical geometry holds key insights into development, aging, disease progression, and cognition. We propose a new framework to study *cortical sheet* geometry to detect inter-group differences and regress against clinical variables. Our model subsumes the complementary properties of folding and thickness.

Typical studies of the cortex focus either on folding [19,25,27] or thickness [9,14,23], but *not* both. They rely on spatial normalization and perform hypothesis tests at every cortical location in the normalized space. However, while the major sulcal patterns are similar across individuals, there is a large variation of the finer-scale cortical folding structure across individuals. This leads to a limited number of homologous features across subject brains [15,16,24], which compromises the precision of the spatial alignment. Moreover, the optimization underlying dense diffeomorphic nonlinear registration is challenged by

multiple local optima. Thus, misregistration is inevitable and can inflate the variability of the cortical descriptors and, thereby, reduce the power of subsequent hypothesis tests. To alleviate problems of partial homology and local optima, we use local histogram-based descriptors that are robust to misregistration.

Our histogram-based descriptor lies on a Riemannian manifold. We propose novel *Riemannian statistical* methods for two typical studies: (i) detecting group differences, using *Mercer kernels* to fit compact models to non-Gaussian data distributions by implicitly employing a lifting map to a reproducing kernel Hilbert space (RKHS), and (ii) regression against clinical variables, using nonlinear regression via *kernel density estimation*. Our kernels exploit the Riemannian structure and enable us to effectively deal with (i) nonlinear variations in the data, i.e., within distributions of cortical descriptors, and (ii) dependencies of the cortical descriptors on clinical variables.

We present a new framework for two typical cortical analyses: (i) detecting group differences and (ii) regression against clinical variables. We use a neighborhood-based histogram feature of local cortical geometry, subsuming folding and thickness characteristics, which is robust to misregistration stemming from biology (partial homology) and numerical analysis (local optima). We propose *regularized nonlinear models* for both analyses, giving compact and reliable model fits, using *kernels* that exploit the Riemannian structure of the space of histograms. Results on simulated and clinical data shows the improved accuracy and stability of our approach cortical sheet analyses.

2 Related Work

Several cortical studies have focus on its thickness characteristics [9,14,23] or folding characteristics [19,25,27], relying on analyses of cortical features across subjects at every location on the cortical surface (typically, the inner/mid cortical surface). Unlike these studies, we model the 3D *cortical sheet* including the complementary characteristics of folding and thickness. There are many models for local cortical folding complexity. While [3] uses local curvature measures, [25] represents the cortical surface using spherical wavelets that lead to a set of wavelet coefficients capturing folding at every cortical location. Another method [26] uses spherical harmonic constructions and local fractal dimension to measure folding complexity. In contrast, we capture folding through curvedness and shape index that can lead to easier interpretation. Unlike other approaches, we perform pointwise studies using a local-histogram descriptor combining folding and thickness, which is robust to misregistration. Local histograms were used in [2] without Riemannian analyses and in [1] for linear Riemannian analyses.

Unlike other descriptors [2,3,25,26], our descriptor lies on a *Riemannian* manifold and we propose novel *kernel-based statistical* methods that exploit the Riemannian structure. Instead of learning a linear approximation to a manifold (e.g., using isomap) and then using Euclidean statistical analyses in the linearized space, we propose nonlinear statistical analysis directly in the (Riemannian) feature space. For group-difference analysis, we propose a *Mercer kernel* for modeling the (possibly non-Gaussian) distributions of the histograms within Riemannian space, along with regularized analysis in RKHS. For

regression, we propose nonlinear *kernel regression* directly on the Riemannian manifold [21], and propose a data-driven consistent estimator of the bandwidth parameter.

3 Methods

We first describe our model for the 3D cortical sheet and the robust descriptor of local cortical geometry. We then describe the proposed Riemannian analyses, involving kernel methods, for hypothesis testing to (i) detect group differences in cortical geometry and (ii) regress cortical geometry against clinical variables.

3.1 A Geometrical Model for the Cerebral Cortical Sheet

We use a medial model for the 3D cortical sheet to represent its folding and thickness characteristics. We choose the mid-cortical surface $\mathcal{M}$ as the medial surface to model folding. In addition, we use pointwise thickness measurements on the medial surface to model the 3D sheet. At each point m on $\mathcal{M}$, the thickness measurement $t(m)$ gives the locations of the inner and outer (pial) cortical surfaces, at distances $t(m)/2$ along the inward and outward normals to $\mathcal{M}$ at m .

We compute cortical thickness based on an improved version of the diffusion-based method in [11], accounting for partial-volume effects by spatially modulating the diffusion strength proportionately. This also produces the mid-cortical surface $\mathcal{M}$ as a level set of the steady-state potential field. For each point m on $\mathcal{M}$, we model the local folding geometry through the principal curvatures determined by the first and second fundamental forms, by fitting a quadratic patch to the local surface around m . The principal curvatures $\kappa_{\min}(m)$ and $\kappa_{\max}(m)$ completely describe the local geometry (up to second order and up to a translation and rotation). We reparametrize the 2D space of $(\kappa_{\min}(m), \kappa_{\max}(m))$, via a polar transformation [13], into the orthogonal bases of curvedness $C(m) := (\kappa_{\min}(m)^2 + \kappa_{\max}(m)^2)^{0.5}$ and shape

index $S(m) := \frac{1}{\pi} \arctan \left(\frac{\kappa_{\min}(m) + \kappa_{\max}(m)}{\kappa_{\min}(m) - \kappa_{\max}(m)} \right)$ that meaningfully separate notions of bending and shape for easier interpretation. The curvedness $C(m) \geq 0$ measures local patch bending at a particular patch scale/size, and is invariant to location and pose. The shape index $S(m) \in [-1, 1]$ is a pure measure of shape, modulo size, location, and pose.

A Robust Local Descriptor of Cortical Folding and Thickness—Nature limits the homology across human brains to about two dozen landmarks in each hemisphere [24]. In typical analyses relying on dense nonlinear diffeomorphic spatial normalization, this partial homology can lead to pointwise correspondences at non-homologous locations, e.g., a gyrus in one brain mapped to a sulcus in another. Such correspondences can inflate the variability of the cortical descriptors and reduce the power of the subsequent hypothesis tests.

One may consider, for a moment, a naive approach to reduce this variability by surface-based spatial smoothing of the descriptors at multiple scales. However, spatial averaging can reduce differences between the smoothed features across corresponding points in subject brains, because multiple spatial distributions of features can all yield the same average. This problem is serious for the brain cortex that is a complex convoluted structure with large variations in the local geometry over small displacements of surface location m . For

instance, the value of the shape index drastically changes magnitude and sign as we traverse a surface path from a gyrus (-1) to a sulcus ($+1$). Curvedness values, along the gyrus-to-sulcus path also changes drastically; from a large positive value to zero (at the inflection point) and back to a large positive value. Cortical thickness also exhibits a variation from gyrus to sulcus, where the crowns of gyri are typically 20% thicker than the fundi of sulci [7]. Indeed, Sect. 4 demonstrates the reduction in the quality of results from smoothing-based *multiscale* features.

Instead of averaging features across spatial neighborhoods, we propose to use the histograms of neighborhood features, which captures the higher-order statistics of the features. In this way, we extend the notion of histogram-based texture descriptors for images to histogram-based descriptors for the cortical sheet geometry; indeed the cortical sheet exhibits high-frequency spatial variations of features just like a textured image exhibits high-frequency spatial variation of intensity. We propose the local cortical descriptor at location m on the cortical surface to be the *joint histogram* of shape-index values $S(n)$, curvedness values $C(n)$, and thickness values $T(n)$ at locations n within a spatial neighborhood around location m . The neighborhood size depends on the typical size of *regions* (*not* individual points) in the cortex over which homologies can be reliably established (we take this as a region of diameter 1 cm). This histogram is robust to the inevitable misregistration of sulci/gyri at fine scales. Moreover, unlike the neighborhood average that is a scalar, the histogram is a richer descriptor that restricts averaging to each histogram bin.

Riemannian Space of Histograms—We perform statistical analyses using the joint histograms $H_i(m)$ as the local feature descriptor for the cortex at location m for subject i . If the histogram has B bins, then $H_i(m) \in (\mathbb{R}_{>0})^B$, $\|H_i(m)\|_1 = 1$, and $H_i(m)$ lies on a Riemannian manifold. To measure distance between two histograms $H_1(m)$ and $H_2(m)$, we use the Fisher-Rao distance metric $d_{\mathcal{H}}(H_1(m), H_2(m)) := d_{\mathcal{F}}(F_1(m), F_2(m))$, where $F_i(m)$ is the square-root histogram that is denoted $\sqrt{H_i(m)}$, with the value in the b -th bin $F_i(m, b) := \sqrt{H_i(m, b)}$, and $d_{\mathcal{F}}(F_1(m), F_2(m))$ is the geodesic distance between $F_1(m)$ and $F_2(m)$ on the unit hypersphere $\mathbb{S}^{B-1}$.

3.2 Group Analysis Using a Mercer Kernel in Riemannian Space

To detect differences between two groups of histograms, we need models for the distribution of histograms within each group. We propose a parametric and a nonparametric model, and associated test statistics for hypothesis testing.

Parametric Model—Modeling a parametric probability density function (PDF) on a hypersphere entails fundamental trade-offs between model generality and the viability of the underlying parameter estimation. For instance, although Fisher-Bingham PDFs on $\mathbb{S}^d$ are able to model generic anisotropic distributions using $\mathcal{O}(d^2)$ parameters, their parameter estimation may be intractable [17]. In contrast, parameter estimation for the $\mathcal{O}(d)$ -parameter von Mises-Fisher (vMF) PDF is tractable, but that PDF can only model isotropic distributions. We use a tractable approximation of a Normal law on a Riemannian manifold [22], modeling anisotropy through its covariance parameter in the tangent space at the mean.

For a group with I subjects, at each cortical location m , we fit the approximate Normal law to the data $\{\sqrt{H_i(m)}\}_{i=1}^I$ as follows. We optimize for the Frechet mean $\mu \in \mathbb{S}^{B-1}$ via iterative gradient descent [4], where $\mu := \arg \min_{\nu} \sum_{i=1}^I d_{\mathcal{H}}^2(\nu, \sqrt{H_i(n)})$ under the constraint $\nu \in \mathbb{S}^{B-1}$. We use the logarithmic map $\text{Log}_{\mu}(\cdot)$ to map the square-root histograms $\{\sqrt{H_i(m)}\}_{i=1}^I$ to the tangent space at the estimated Frechet mean μ and find the optimal covariance matrix C in closed form [8]. For any histogram H , we define the geodesic Mahalanobis distance between $\sqrt{H}$ and mean μ , given covariance C , as

$$d_{\mathcal{H}}(\sqrt{H}; \mu, C) := (\text{Log}_{\mu}(\sqrt{H})^{\top} C^{-1} \text{Log}_{\mu}(\sqrt{H}))^{0.5}.$$

Nonparametric Model—Our nonparametric modeling approach relies on the use of a Mercer kernel to map the square-root histograms on the hypersphere to a RKHS, followed by using (kernel) PCA to fit a multivariate Gaussian distribution to the mapped data in the RKHS. In this way, the Mercer kernel allows us to model nonlinear modes of variation in the data more effectively than the Normal-law approach. The kernel-based strategy has the potential to yield compact fits to the data, thereby being able to detect subtle differences, and detect cortical regions, more accurately. For this purpose, we choose a real-valued Mercer kernel that exploits the Riemannian structure of the space of histograms, which is the vMF Mercer kernel defined on the Riemannian space of histograms as follows:

$K_{\text{vMF}}(H_1, H_2; \gamma) := \exp(\gamma \langle \sqrt{H_1}, \sqrt{H_2} \rangle)$, where $\gamma > 0$ is a regularization parameter. We set γ to be the average of all intra-group pairwise inner products between the square-root histograms in the dataset.

Let the associated lifting map for the Mercer kernel be $\Phi(\cdot)$ that maps H to $\Phi(H) := k(\cdot, H)$ in a RKHS $\mathcal{F}$. Consider two vectors in RKHS: $f := \sum_{i=1}^I \alpha_i \Phi(H_i)$ and $f' := \sum_{j=1}^J \beta_j \Phi(H_j)$.

The inner product $\langle f, f' \rangle_{\mathcal{F}} := \sum_{i=1}^I \sum_{j=1}^J \alpha_i \beta_j K_{\text{vMF}}(H_i, H_j)$. The norm

$\|f\|_{\mathcal{F}} := \sqrt{\langle f, f \rangle_{\mathcal{F}}}$. When $f, f' \in \mathcal{F} \setminus \{0\}$, let $f \otimes f'$ be the rank-one operator defined as $f \otimes f'(g) := \langle f', g \rangle_{\mathcal{F}} f$. The model fit via kernel PCA, to a sample of size N , gives a mean μ and a covariance operator C . The Mahalanobis distance relies on a regularized sample inverse-

covariance operator [18] $C^{-1} := \sum_{q=1}^Q (1/\lambda_q) v_q \otimes v_q$, where λ_q is the q -th largest eigenvalue of the covariance operator C , v_q is the corresponding eigenfunction, and $Q < N$ is a regularization parameter. We choose Q to be the number of principal eigenfunctions that capture 90% of the energy in the eigenspectrum. Then, the square-root inverse-covariance operator is $C^{-1/2} := \sum_q (1/\sqrt{\lambda_q}) v_q \otimes v_q$ and the Mahalanobis distance of point $f \in \mathcal{F}$ from mean $\mu \in \mathcal{F}$ is $d_{\mathcal{F}}(f, \mu, C) := \|C^{-1/2}(f - \mu)\|_{\mathcal{F}}$.

Test Statistics—For both parametric and nonparametric models, at each cortical location m , for two cohorts $\{H_i^X(m)\}_{i=1}^I$ and $\{H_j^Y(m)\}_{j=1}^J$, we obtain the means $\mu^X(m)$, $\mu^Y(m)$ and covariances $C^X(m)$, $C^Y(m)$. For the parametric model the means lie on the hypersphere and

the covariances are defined on the tangent spaces at the means. For the nonparametric model, the means and the covariance operators lie in the RKHS associated with the Mercer kernel. To perform hypothesis testing for significant differences between the groups, we need test statistics for the parametric and nonparametric models. For both models, we propose a test statistic $t(m)$ to measure the dissimilarity between the two cohort distributions by adding the squared Mahalanobis geodesic distance between the group means with respect to each group covariance, i.e.,

$t(m) := d_{\mathcal{M}}^2(\mu^X(m); \mu^Y(m), C^Y(m)) + d_{\mathcal{M}}^2(\mu^Y(m); \mu^X(m), C^X(m))$. The Hotelling's 2-sample T^2 test statistic used in the usual Euclidean multivariate Gaussian case may be inapplicable in our parametric Riemannian case because the covariances $C^X(m)$ and $C^Y(m)$ are defined in two different (tangent) spaces.

3.3 Kernel Regression in Riemannian Space

To regress local cortical geometry against a clinical variable, we need a regressor in the Riemannian space of histograms. We propose a nonlinear regressor that relies on principles on kernel density estimation.

Regression Model—We propose an extension of the Nadaraya-Watson nonlinear kernel regressor [10] to the Riemannian space of histograms [21], with the kernel PDF as $PK(H; \mu, c) := (1/\eta) \exp(cd_{\mathcal{H}}(H, \mu))$, where H and μ are histograms, μ is the mean parameter, c is the concentration parameter, and η is the normalization constant. Interestingly, this kernel PDF is equivalent to the vMF PDF on the hypersphere associated with square-root histograms.

Given a cohort comprising the independent variables as the joint histograms $\{H_i(m)\}_{i=1}^I$ and the dependent variable as the clinical score $\{S_i\}_{i=1}^I$ the clinical score S for an arbitrary

histogram H is given by
$$r(H; \{H_i, S_i\}_{i=1}^I, c) := \frac{\sum_{i=1}^I P_K(H; H_i, c) S_i}{\sum_{i=1}^I P_K(H; H_i, c)}.$$

Data-driven selection of the bandwidth parameter $1/c$ is critical for reliable regression. We propose a leave-one-out (LOO) cross validation approach [10] to automatically optimize the bandwidth parameter based on the structure underlying the data. We optimize c by brute force; we first compute the mean squared error, in the LOO regression estimates across the dataset, for a sequence of discretized parameter values and then select the parameter value that minimizes this error. This estimate c_{CV} is guaranteed to be consistent [10].

Test Statistic—At each location m on the cortex, we want to test the hypothesis that the clinical variable is dependent on the local cortical descriptor. If the null hypothesis was true, i.e., the clinical variable was independent of the local cortical geometry, then the optimal regressor would be a constant function. We compare the sets of residual errors across the

dataset $\{S_j - \hat{r}(H_j(m); \{H_i(m), S_i\}_{i=1}^I)\}_{j=1}^I$, resulting under two regressors $\hat{r}(\cdot)$, i.e., (i) the constant-function regressor $\hat{r}(H) = \delta$ and (ii) the kernel regressor

$\hat{r}(H) = r(H; \{H_i(m), S_i\}_{i=1}^I, c_{CV})$. If the null hypothesis was true, then the two regressors will produce residuals that would appear to be drawn from the same distribution. If the null hypothesis was false, then the residuals from the kernel regressor would be significantly

smaller than those from the constant-function regressor. Thus, we propose the test statistic as the well-known F statistic, i.e., the ratio of the sample variances of the two sets of residuals.

3.4 Permutation Testing and Stability Analysis Using Resampling

Voxel-wise parametric hypothesis testing in the framework of general linear models runs a test at each voxel and adjusts p -values to control for Type-I error arising from multiple comparisons, using Gaussian field theory. However, such parametric approaches make strong assumptions on the data distributions and the dependencies within neighborhoods. In contrast, permutation tests are nonparametric, rely on the generic assumption of exchangeability, lead to stronger control over Type-I error, are more robust to deviations of the data and effects of processing from an assumed model, and yield multiple-comparison adjusted p values (controlling family-wise error rate through the permutation-distribution of the extremal test statistic). Under the null hypothesis, assuming exchangeability, in mappings between (i) cortical descriptors and (ii) group labels/clinical scores, we use permutation-based resampling to empirically compute the null distribution of the extremal (across all cortical locations) test-statistic values proposed in Sects. 3.2 and 3.3, to compute p values. To evaluate the *stability* of the p values under variation in cohorts, we compute a set of p values by bootstrap sampling the original cohort and analyzing the associated variation.

4 Results and Discussion

We evaluate our framework on MRI volumes from 2 large datasets: (i) OASIS, for the group-difference study, and (ii) Human Connectome Project (HCP), for the regression study. We use BrainSuite (<http://brainsuite.org/>) for computing the tissue segmentation, thickness (Fig. 1(a)), mid-cortical surface, shape index (Fig. 1(b)), curvedness (Fig. 1(c)), and spatial normalization [12].

Group-Difference Study: Validation with Simulated Differences

We randomly assigned 140 control subjects to 2 groups of 90 and 50 subjects. We treat the larger group as normal, mimicking practical scenarios. For the 50 subjects, we simulated both cortical thinning (eroding the cortex segmentation) and flattening (smoothing the cortex segmentation) in a part of the right parietal lobe (Fig. 1(d)). This (i) reduced thickness and curvedness values and (ii) increased the concentration of shape index values around ± 0.5 (corresponding to gyral ridges and sulcal valleys). We then tested for differences between cortices of these 2 cohorts. We compare our approach, i.e., using local joint-histogram descriptors and *nonparametric kernel-based* Riemannian statistical analysis (Sect. 3.2), against 3 other approaches: (i) *multiscale* joint descriptors of shape index, curvedness, and thickness (Sect. 3.1), with Euclidean analysis, (ii) histogram descriptors with Euclidean multivariate-Gaussian analysis, ignoring the Riemannian structure, and (iii) histogram descriptors with Riemannian *parametric Normal-law* modeling (Sect. 3.2).

The detection of group differences with multiscale descriptors (Fig. 2(a)) leads to many false positives. While the histogram-based Euclidean analysis (Fig. 2(b)) detects a region much smaller than the actual (false negatives), the histogram-based Riemannian Normal-law

analysis detects a region larger than the actual (false positives). Detection accuracy is best with the proposed approach of Riemannian analysis with Mercer kernels (Fig. 2(d)) that leads to a compact region with very few false positives/negatives. The stability of all histogram-based analyses (Fig. 3) is significantly higher (lower variation in p values) than the multiscale-descriptor analysis. The stability of the proposed approach (Fig. 3(d)), in the affected region in the parietal lobe, is higher than the other histogram-based approaches (Fig. 2(b) and (c)).

Group-Difference Study: MCI Data

We tested for differences in 2 cohorts from the OASIS dataset: (i) 140 control subjects and (ii) 28 subjects with mild cognitive impairment (MCI) with a clinical dementia rating of 1. The histogram-based descriptors (Fig. 5(b) and (c)) give more stable results than the multiscale descriptors (Fig. 5(a)). The proposed kernel-based approach (Fig. 4(c)) is able to detect differences in more brain lobes, without significant reduction in stability, as compared to the histogram-based Riemannian Normal-law approach (Fig. 4(b)).

Regression Study: Validation with Simulated Effects

We took 200 MRI volumes from the HCP dataset along with clinical scores on a task. First, we randomly reassign the scores to subjects, to eliminate any existing dependencies between the score and cortical geometry. Then, for all the subjects, we simulated subtle relationships between the scores and the geometry of a part of the right parietal lobe (Fig. 1(d)), by cortical thinning (eroding the cortex segmentation) and flattening (smoothing the cortex segmentation) in amounts directly related to the score.

We compare our approach, i.e., using local joint-histogram descriptors and nonlinear kernel regression on Riemannian space (Sect. 3.3) with cross-validation based data-driven bandwidth estimation, against 3 other approaches: (i) multiscale joint descriptors (Sect. 3.1) with linear regression, (ii) multiscale joint descriptors with nonlinear regression with data-driven bandwidth estimation, and (iii) histogram descriptors with linear regression, ignoring Riemannian structure.

Linear regression with multiscale descriptors performs poorly, wrongly indicating relationships (Fig. 6(a)) with the entire cortex and with the worst stability under bootstrap sampling of the cohort (Fig. 7(a)). With multiscale descriptors, nonlinear regression improves over linear regression, but still leads to many false positive and false negatives (Fig. 6(b)) and has less stability (Fig. 7(b)) than histogram-based descriptors. Histogram-based descriptors perform much better than multiscale descriptors, yielding a coherent region with few false positives (Fig. 6(c) and (d)) and greater stability (Fig. 7(c) and (d)). Our histogram-based nonlinear-regression on Riemannian space, along with automatic bandwidth selection, yields the most stable results (Fig. 7(d)). Our data-driven bandwidth-parameter estimation, at every cortical location, is important in adding to the stability of the nonlinear regression approaches.

Regression: Task Score, Stability Analysis

We tested for dependencies between local cortical structure and the task score (average of language-based story and mathematics tasks) on 200 subjects in the HCP dataset. Analogous to the other results in the paper, the histogram-based descriptors (Fig. 9) give far more stable results. The proposed histogram-based approaches, with linear regression ignoring Riemannian structure (Fig. 8(a) and (b)) and nonlinear kernel regression in Riemannian space (Fig. 8(c) and (d)), seem to show dependencies of the task scores on the cortical structure around the insular and opercular regions and near language areas. Recent studies on language processing, using functional imaging, in the brain seem to indicate insula to play an important role in language processing [5,6,20]. The insula shares reciprocal functional and structural connections with linguistic, motor, limbic, and sensory brain areas.

Conclusion

We propose a framework for analysis of cortical geometry to (i) detect group differences and (ii) regress against clinical variables. We use a cortical descriptor that is *robust* to fine-scale misregistration, having a *Riemannian* structure. We propose novel *regularized nonlinear models* for analyses, giving compact and reliable model fits, using kernels that exploit the Riemannian structure. Our results show improved robustness and accuracy relative to (i) multiscale descriptors and (ii) linear analyses.

Acknowledgments

Thanks to IIT-B Seed Grant 14IRCCSG010, NIH R01NS089212, NIH R01NS074980.

References

1. Awate, SP., Leahy, RM., Joshi, AA. Riemannian statistical analysis of cortical geometry with robustness to partial homology and misalignment. In: Ourselin, S., Joskowicz, L., Sabuncu, MR., Unal, G., Wells, W., editors. MICCAI 2016. LNCS. Vol. 9900. Springer; Cham: 2016. p. 237-246.
2. Awate SP, Yushkevich P, Song Z, Licht D, Gee J. Cerebral cortical folding analysis with multivariate modeling and testing: studies on gender differences and neonatal development. *Neuroimage*. 2010; 53(2):450–459. [PubMed: 20630489]
3. Batchelor P, Castellano-Smith A, Hill D, Hawkes D, Cox T, Dean A. Measures of folding applied to the development of the human fetal brain. *IEEE Trans. Med. Imaging*. 2002; 21(8):953–965. [PubMed: 12472268]
4. Buss S, Fillmore J. Spherical averages and applications to spherical splines and interpolation. *ACM Trans. Graph*. 2001; 20(2):95–126.
5. Chee M, Soon C, Lee H, Pallier C. Left insula activation: a marker for language attainment in bilinguals. *Proc. Natl. Acad. Sci*. 2004; 101(42):15265–15270. [PubMed: 15469927]
6. Eickhoff S, Heim S, Zilles K, Amunts K. A systems perspective on the effective connectivity of overt speech production. *Phil. Trans. Royal Soc*. 2009; 367:2399.
7. Fischl B, Dale A. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc. Natl. Acad. Sci*. 2000; 97(20):11050–11055. [PubMed: 10984517]
8. Fletcher T, Lu C, Pizer S, Joshi S. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imaging*. 2004; 23(8):995–1005. [PubMed: 15338733]
9. Hardan A, Muddasani S, Vemulapalli M, Keshavan M, Minshew N. An MRI study of increased cortical thickness in autism. *Am. J. Psychiatry*. 2006; 163(7):1290–1292. [PubMed: 16816240]
10. Hardle, W. *Applied Nonparametric Regression*. Cambridge University Press; Cambridge: 1990.

11. Jones S, Buchbinder B, Aharon I. Three-dimensional mapping of cortical thickness using Laplace's equation. *Hum. Brain Mapp.* 2000; 11(1):12–32. [PubMed: 10997850]
12. Joshi, AA., Shattuck, DW., Leahy, RM. A method for automated cortical surface registration and labeling. In: Dawant, BM.Christensen, GE.Fitzpatrick, JM., Rueckert, D., editors. *WBIR 2012. LNCS. Vol. 7359.* Springer; Heidelberg: 2012. p. 180-189.
13. Koenderink, JJ. *Solid Shape.* MIT Press; Chichester: 1991.
14. Luders E, Narr K, Thompson P, Rex D, Woods R, Jancke L, Toga A. Gender effects on cortical thickness and the influence of scaling. *Hum. Brain Mapp.* 2006; 27:314–324. [PubMed: 16124013]
15. Lyttelton O, Boucher M, Robbins S, Evans A. An unbiased iterative group registration template for cortical surface analysis. *Neuroimage.* 2007; 34:1535–1544. [PubMed: 17188895]
16. Mangin J, Riviere D, Cachia A, Duchesnay E, Cointepas Y, Papadopoulos-Orfanos D, Scifo P, Ochiai T, Brunelle F, Regis J. A framework to study the cortical folding patterns. *Neuroimage.* 2004; 23(1):S129–S138. [PubMed: 15501082]
17. Mardia, K., Jupp, P. *Directional Statistics.* Wiley; Hoboken: 2000.
18. Mas A. Weak convergence in the function autoregressive model. *J. Multivar. Anal.* 2007; 98:1231–1261.
19. Nordahl C, Dierker D, Mostafavi I, Schumann C, Rivera S, Amaral D, Van-Essen D. Cortical folding abnormalities in autism revealed by surface-based morphometry. *J. Neurosci.* 2007; 27(43):11725–11735. [PubMed: 17959814]
20. Oh A, Duerden E, Pang E. The role of the insula in speech and language processing. *Brain Lang.* 2014; 135:96–103. [PubMed: 25016092]
21. Pelletier B. Nonparametric regression estimation on closed Riemannian manifolds. *J. Nonparametric Stat.* 2006; 18:57–67.
22. Pennec X. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *J. Math. Imaging Vis.* 2006; 25(1):127–154.
23. Redolfi A, Manset D, Barkhof F, Wahlund L, Glatard T, Mangin JF, Frisoni G. Head-to-head comparison of two popular cortical thickness extraction algorithms: a cross-sectional and longitudinal study. *Plos ONE.* 2015; 10(3):e0117692. [PubMed: 25781983]
24. Van-Essen D, Dierker D. Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron.* 2007; 56:209–225. [PubMed: 17964241]
25. Yeo, BTT., Yu, P., Grant, PE., Fischl, B., Golland, P. Shape analysis with over-complete spherical wavelets. In: Metaxas, D.Axel, L.Fichtinger, G., Székely, G., editors. *MICCAI 2008. LNCS. Vol. 5241.* Springer; Heidelberg: 2008. p. 468-476.
26. Yotter R, Nenadic I, Ziegler G, Thompson P, Gaser C. Local cortical surface complexity maps from spherical harmonic reconstructions. *Neuroimage.* 2011; 56:961. [PubMed: 21315159]
27. Yu P, Grant P, Qi Y, Han X, Segonne F, Pienaar R, Busa E, Pacheco J, Makris N, Buckner R, Golland P, Fischl B. Cortical surface shape analysis based on spherical wavelets. *IEEE Trans. Med. Imaging.* 2007; 26(4):582–597. [PubMed: 17427744]

Fig. 1. Cortical sheet model. Computed values of the cortical (a) thickness, (b) shape index, and (c) curvedness, at each point on the mid-cortical surface. Simulating cortical effects. (d) Selected region for simulating group effects.

Fig. 2.

Group-difference study: validation with simulated differences. permutation test p values with the following features: (a) joint multiscale descriptor, (b) joint histogram descriptor with Euclidean analysis, (c) joint histogram descriptor with Riemannian normal-law analysis, (d) joint histogram descriptor with Riemannian kernel-based analysis (*proposed*).

Fig. 3.

Group-difference study: simulated differences, stability analysis. *Standard deviation* of p values, via *bootstrap sampling*, for: (a) joint multiscale descriptor, (b) joint histogram descriptor with Euclidean analysis, (c) joint histogram descriptor with Riemannian normal-law analysis, (d) joint histogram descriptor with Riemannian kernel-based analysis (*proposed*).

Fig. 4. Group-difference study: MCI. Permutation test p values with: (a1), (a2) joint multiscale descriptor, (b1), (b2) joint histogram descriptor with Riemannian Normal-law analysis, (c1), (c2) joint histogram descriptor with Riemannian kernel-based analysis (*proposed*).

Fig. 5. Group-difference study: MCI, stability analysis. *Standard deviation* of permutation test p values, via *bootstrap sampling*, for: (a1), (a2) joint multiscale descriptor, (b1), (b2) joint histogram descriptor with Riemannian normal-law analysis, (c1), (c2) joint histogram with Riemannian kernel-based analysis (*proposed*).

Fig. 6.

Regression study: validation with simulated effects. Permutation test p values for: (a) joint multiscale descriptor with linear regression, (b) joint multiscale descriptor with kernel regression, (c) joint histogram descriptor with linear regression, (d) joint histogram descriptor with kernel regression in Riemannian space (*proposed*).

Fig. 7.

Regression study: simulated effects, stability analysis. *Standard deviation* of p values, via *bootstrap sampling*, for: (a) joint multiscale descriptor with linear regression, (b) joint multiscale descriptor with kernel regression, (c) joint histogram descriptor with linear regression, (d) joint histogram descriptor with kernel regression in Riemannian space (*proposed*).

Fig. 8.

Regression study: task score. Permutation test p values for the joint histogram descriptor with: (a) and (b) linear regression, and (c) and (d) kernel regression in Riemannian space, with data-driven bandwidth selection (*proposed*).

Fig. 9.

Regression study: task score, stability analysis. *Standard deviation* of permutation test p values, via *bootstrap sampling*, for the: (a) joint multiscale descriptor with linear regression, (b) joint multiscale descriptor with kernel regression, (c) and (d) joint histogram descriptor with kernel regression in Riemannian space, with data-driven bandwidth selection (*proposed*).