CORTICAL CORRESPONDENCE USING ENTROPY-BASED PARTICLE SYSTEMS AND LOCAL FEATURES

Ipek Oguz[1], *Joshua Cates*[3], *Thomas Fletcher*[3], *Ross Whitaker*[3], *Derek Cool*[4], *Stephen Aylward*[5], *Martin Styner*[1,2]

Departments of [1]Computer Science and [2]Psychiatry, University of North Carolina, Chapel Hill, USA
[3]School of Computing, University of Utah, Salt Lake City UT, USA
[4]Robarts Research Institute, London, ON, Canada [5]Kitware Inc., Clifton Park NY USA

ABSTRACT

This paper presents a new method of constructing compact statistical point-based models of populations of human cortical surfaces with functions of spatial locations driving the correspondence optimization. The proposed method is to establish a tradeoff between an even sampling of the surfaces (a low surface entropy) and the similarity of corresponding points across the population (a low ensemble entropy). The similarity metric, however, isn't constrained to be just spatial proximity, but can be any function of spatial location, thus allowing the integration of local cortical geometry as well as DTI connectivity maps and vasculature information from MRA images. This method does not require a spherical parameterization or fine tuning of parameters. Experimental results are also presented, showing lower local variability for both sulcal depth and cortical thickness measurements, compared to other commonly used methods such as FreeSurfer.

Index Terms— Correspondence, Image Shape Analysis, Brain Modeling, Statistics, Image Registration.

1. INTRODUCTION

Statistical modeling of anatomical objects is becoming increasingly important in the segmentation, analysis and interpretation of medical datasets. Constructing such statistical models requires the ability to compute local shape differences among similar objects. This introduces the problem of finding corresponding points. Consistent computation of corresponding points on 3D anatomical surfaces is a difficult task, since manually choosing landmark points not only is cumbersome, but also does not yield a satisfyingly dense correspondence map. It should also be noted that no generic "ground truth" definition of dense correspondence exists across different anatomical surfaces. The choice of particular correspondence metric must, therefore, be application-driven.

In this work, we are introducing a framework for finding corresponding points on populations of human cortical surfaces. The correspondence computation on the cortex is an even more challenging problem due to the highly convoluted geometry of the brain and the high variability of folding patterns across subjects. Using mere spatial locations of surface points produces a weak and inadequate correspondence map. Our work allows the usage of more local information, called 'features' throughout this manuscript, for computing correspondence. We demonstrate our technique using a sulcal depth function as an additional feature; the framework also allows for more complex features, such as connectivity maps computed from DTI images, or vessel structure extracted from MRA images, to be used for establishing correspondence. The particular choice of features should be determined by the target applications.

We are using a particle based entropy minimizing system, as introduced by Cates et al.[1]. We extend the particle framework to allow the local features to drive the correspondence optimization. We also use a cortex flattening technique to overcome difficulties originating from the highly convoluted nature of the cortical surface. The high level processing pipeline, therefore, consists of 'inflating' the cortical surfaces, computing correspondence on the inflated surface, and 'deflating' to obtain corresponding points on the original brain surface. We evaluate our results based on how well they reduce local variability on both sulcal depth and cortical thickness datasets.

2. PREVIOUS WORK

Various automated methods for establishing correspondence have been suggested. One of the earliest such methods was proposed by Kotcheff and Taylor in [2], where they tried minimizing information content across an ensemble with a cost function $\sum_k \log(\lambda_k + \alpha)$, where λ_k are the eigenvalues of the covariance matrix and α is a regularization term. Davies et al. [3, 4] propose that the simplest description of a population is the best; simplicity is measured in terms of the length of the code to transmit the data as well as the model parameters, hence the name of their method, Minimum Description Length (MDL). MDL implementations in 3D usually rely on spherical parameterizations of the surfaces, which must be obtained through a preprocessing step such as proposed in [5] that relaxes a spherical parameterization onto the input mesh. In [6], Heimann et al. present a gradient descent optimization method for the MDL algorithm and explore using local curvature in addition to spatial locations in the MDL cost function.

Styner et al.[7] describe an empirical study that shows ensemble-based statistics improve correspondences relative to pure geometric regularization, and that MDL performance is virtually the same as that of min-log $|\Sigma + \alpha I|$. This last observation is consistent with the well-known result from information theory: MDL is, in general, equivalent to minimum entropy [8]. Cates et al. [1, 9] propose a system exploring this property; this is the underlying technique for

This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. Information on the National Centers for Biomedical Computing can be obtained from http://nihroadmap.nih.gov/bioinformatics. This research is also partially funded by the NIH/NCRR Center for Integrative Biomedical Computing, P41-RR12553-07, and UNC Neurodevelopmental Disorders Research Center, HD 03110.

the methodology presented in this paper and will be discussed in more detail in Sec. 3.2.

The FreeSurfer system [10, 11, 12, 13, 14, 15] provides an entire framework for the segmentation, surface reconstruction, topology correction, inflation and spherical parameterization of the cortex. The correspondence across population of the cortical surfaces is established through a semi-rigid alignment of the folding patterns on each cortex to an average cortex shape, using the spherical parameterizations. The folding pattern is quantized by the local sulcal depth values. Note that this is essentially a two-step correspondence computation, in that it first optimizes the spatial correspondence (while computing the spherical parameterization) and then optimizes the sulcal depth correspondence. This method also is different from ours in that it focuses on a pair-wise correspondence (subject to average), whereas our method emphasizes a group-wise approach.

3. METHODOLOGY

In this work, we are presenting a cortical correspondence system that incorporates various functions of spatial coordinates. We choose to use a particle based entropy minimizing system, as introduced by Cates et al [1, 9], for the correspondence computation in a population-based manner. We extend the methodology presented in [1, 9] to allow the use of local functions for cortical correspondence. This is critical in obtaining an application-specific correspondence, since various local measures such as sulcal depth, cortical thickness, or DTI measurements can be used depending on the particular clinical context. Furthermore, given the highly folded and curved nature of the cortex surface, distances measured in 3D space between points does not reflect the actual distance along the cortical sheet (e.g. in the case of two points lying on different banks of a sulcus); it therefore makes little sense to use spatial proximity as a standalone measure of correspondence strength.

The main idea for the entropy based correspondence method is to construct a point-based sampling of the shape ensemble that simultaneously maximizes both the geometric accuracy and the statistical simplicity of the model. Surface point samples, which also define the shape-to-shape correspondences, are modeled as sets of dynamic particles that are constrained to lie on a set of implicit surfaces. Sample positions are optimized by gradient descent on an energy function that balances the negative entropy of the distribution on each shape with the positive entropy of the ensemble of shapes. Local measurements on the cortical surface, as will be discussed in Sec. 3.2.3, are incorporated into this ensemble entropy to provide a general correspondence definition.

One of the main challenges of using the particle based entropy minimizing technique on the cortical surface is that it assumes the particles to be existing on local tangent planes, which presents a problem for the cortex given the highly convoluted surface geometry. We overcome this difficulty by first 'inflating' the cortex surface. This way, we obtain a less convoluted, blob-like surface, for the particles to interact on. However, we need a one-to-one correspondence between the original cortex surface and the inflated surface, since the data to be used for correspondence, such as the sulcal depth, lives on the original cortex surface. A set of automated tools distributed as part of the FreeSurfer [10, 11, 12, 13, 14, 15] package are used to preprocess the data as described in Sec. 3.1.

In order to evaluate the quality of our results, we analyze the local variability of both the features that are being used for the correspondence computation (sulcal depth in this case), and, more importantly, of a different local measurement, namely, cortical thickness.

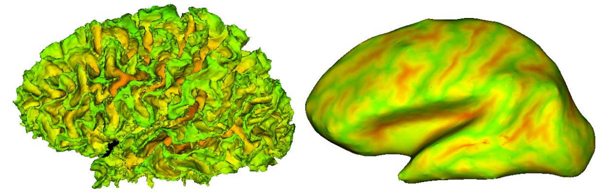

Fig. 1. The sulcal depth pictured as a color map on the white matter surface and its inflation surface. The green and red regions correspond to highly negative and highly positive values of sulcal depth, respectively.

3.1. Surface Inflation, Sulcal Depth Computation, and Other Preprocessing Steps

In this work, we use FreeSurfer for the cortical surface reconstruction as well as surface inflation. The inflation algorithm of FreeSurfer [12] provides a surface representation that is much smoother than the original convoluted cortical surface while minimizing metric distortions. This is achieved via the optimization of an energy functional consisting of the weighted sum of a spring force that works towards 'inflating' the surface and a metric preservation term that ensures that as little metric distortion as possible is introduced in the process. The inflation process is such that points that lie in convex regions move inwards while points in concave regions move outwards over time. Therefore, the average convexity/concavity of the surface over a region, also referred to as sulcal depth, can be computed as the integral of the normal movement of a point during inflation. Specifically, the sulcal depth $C(x_k^0)$ at position x_k is defined as:

$$C(x_k^0) = \int v_t^k \bullet n(k) dt$$

where $n(k)$ is the unit normal vector at position x_k, and v_t^k is the direction in which the vertex x_k moves at time step t, which is equivalent to the partial derivative of the energy functional driving the inflation. It should be noted that C captures the high level foldings of the cortical surface, but is relatively insensitive to the smaller folds; this property makes C an attractive correspondence metric, since it is relatively stable across individuals. Fig. 1 shows the sulcal depth as a color map on both the original white matter surface and the inflated surface.

Once surface construction and inflation is completed, the resulting surfaces are sorted into an octree structure, which provides a fast lookup table for locating the particles in the correspondence optimization. This lookup table is used in the computation of barycentric coordinates on both the original and inflated surfaces, which allows the particles to be efficiently and accurately pulled back and forth between the original white matter surface and the inflated surface.

3.2. Particle Method

3.2.1. Entropy-Based Surface Sampling

In this work, as presented in [9], we sample a surface $\mathcal{S} \subset \Re^d$ using a discrete set of N points that are considered random variables $Z = (X_1, X_2, \ldots, X_N)$ drawn from a probability density function (PDF), $p(X)$. We denote a realization of this PDF with lower case, and thus we have $z = (x_1, x_2, \ldots, x_N)$, where $z \in \mathcal{S}^N$. The probability of a realization x is $p(X = x)$, which we denote simply as

$p(x)$. The amount of information contained in such a random sampling is, in the limit, the differential entropy of the PDF, which is $H[X] = -\int_S p(x) \log p(x) dx = -E\{\log p(X)\}$, where $E\{\cdot\}$ is the expectation. The optimization problem is given by:

$$\hat{z} = \arg\min_z E(z) \text{ s.t. } x_1, \ldots, x_N \in \mathcal{S}. \tag{1}$$

Cates et al. describe in [9] a method for estimating the probability distributions and for minimizing the cost function C, which is an approximation of the (negative) entropy E. The main idea is that the surface points (which are called *particles*) must move away from each other, and there is a set of particles moving under a repulsive force and constrained to lie on the surface. The motion of each particle is away from all of the other particles, but the forces are weighted by a Gaussian function of inter-particle distance. Interactions are therefore local for sufficiently small σ.

The preceding minimization produces a *uniform* sampling of a surface. For some applications, a strategy that samples adaptively in response to higher order shape information is more effective. Cates et al. [9] also provides a detailed description of how to modify the probability estimations to obtain such an adaptive sampling scheme.

Note that this particle formulation computes Euclidean distance between particles, rather than the geodesic distance on the surface. Thus, a sufficiently dense sampling is assumed, so that nearby particles lie in the tangent planes of the zero sets of a scalar function F which provides the implicit cortical surface. This is an important consideration; in cases where this assumption is not valid, such as highly convoluted surfaces, the distribution of particles may be affected by neighbors that are outside of the true manifold neighborhood. The cortex is a prime example of such a highly convoluted surface. In this work, we overcome this problem by inflating the cortical surface prior to optimizing correspondence. The particles therefore live in the tangent planes of the inflated surface; they are only pulled back to the original cortical surface for correspondence evaluation purposes.

3.2.2. Ensemble Entropy of Correspondence Positions

An ensemble $\mathcal{E}$ is a collection of M surfaces, each with their own set of particles, i.e. $\mathcal{E} = z^1, \ldots, z^M$. The ordering of the particles on each shape implies a correspondence among shapes, and thus we have a matrix of particle positions $P = x_j^k$, with particle positions along the rows and shapes across the columns. We model $z^k \in \Re^{Nd}$ as an instance of a random variable Z, and propose to minimize the combined ensemble and shape cost function

$$Q = H(Z) - \sum_k H(P^k), \tag{2}$$

which favors a compact ensemble representation balanced against a uniform distribution of particles on each surface. The different entropies are commensurate so there is no need for ad-hoc weighting of the two function terms.

For this discussion we assume that the complexity of each shape is greater than the number of examples, and so we would normally choose $N > M$. Given the low number of examples relative to the dimensionality of the space, we must impose some conditions in order to perform the density estimation. For this work we assume a normal distribution and model $p(Z)$ parametrically using a Gaussian with covariance Σ. The entropy is then given by

$$H(Z) \approx \frac{1}{2} \log |\Sigma| = \frac{1}{2} \sum_{j=1}^{Nd} \log \lambda_j, \tag{3}$$

where $\lambda_1, ..., \lambda_{Nd}$ are the eigenvalues of Σ.

In practice, Σ will not have full rank, in which case the entropy is not finite. We must therefore regularize the problem with the addition of a diagonal matrix αI to introduce a lower bound on the eigenvalues. We estimate the covariance from the data, letting Y denote the matrix of points minus the sample mean for the ensemble, which gives $\Sigma = (1/(M-1))YY^T$. Because $N > M$, we perform the computations on the dual space (dimension M), knowing that the determinant is the same up to a constant factor of α. Thus, we have the cost function G associated with the ensemble entropy:

$$\log |\Sigma| \approx G(P) = \log \left| \frac{1}{M-1} Y^T Y, \right| \text{and} - \frac{\partial G}{\partial P} = Y(Y^T Y + \alpha I)^{-1}. \tag{4}$$

We now see that α is a regularization on the inverse of $Y^T Y$ to account for the possibility of a diminishing determinant. The negative gradient $-\partial G / \partial P$ gives a vector of updates for the entire system, which is recomputed once per system update. This term is added to the shape-based updates described in the previous section to give the update of each particle:

$$x_j^k \leftarrow \gamma \left[-\partial G / \partial x_j^k + \partial E^k / \partial x_j^k \right]. \tag{5}$$

3.2.3. Ensemble Entropy of Functions of Correspondence Positions

In the case of computing entropy of vector-valued functions of the correspondence positions P, we now consider the more general case where $\tilde{P} = f(x_j^k)$, where $f : \Re^d \to \Re^q$. $\tilde{Y}$ becomes a matrix of the function values at the particle points minus the means of those functions at the points, and we compute the general cost function simply as

$$\tilde{G}(\tilde{P}) = \log \left| \frac{1}{M-1} \tilde{Y}^T \tilde{Y}, \right|. \tag{6}$$

Let $Q = (\tilde{Y}^T \tilde{Y} + \alpha I)^{-1}$ By the chain rule, the partial derivative of $\tilde{G}$ with respect to each shape k becomes

$$-\frac{\partial \tilde{G}}{\partial \tilde{P^k}} = J_k^T Q^k, \tag{7}$$

where J_k is the Jacobian of the functional data for shape k. Note that each J_k is in the form of a block diagonal matrix with $(qxN)x(qxN)$ blocks, with diagonal blocks the qxd submatrices of the function gradients at particle j.

4. RESULTS

We evaluate our method on two different datasets, based on how well the resulting correspondence maps reduce variability of sulcal depth and cortical thickness measurements. Our population consists of 10 cortical surfaces from healthy patients reconstructed via FreeSurfer from T1 images; 1 had to be removed due to failure of succesful surface reconstruction with FreeSurfer, caused by extreme noise in the MRI scan (therefore, the final study used $M = 9$ subjects). All reconstructions showed some degree of error in the temporal lobe; no manual interventions were made. Due to time constraints, refined, manually corrected surfaces will be used only in our future evaluations.

For both the sulcal depth and cortical thickness measurements, we compute the local sample variance of the measurement before and after correspondence optimization; the desirable result is a lowering of variability across corresponding points. We compare variances based on our method's correspondence maps to initial data

	Sulcal Depth	Cortical Thickness
Initial Data	0.227634	0.334858
XYZ-entropy	0.219627	0.341715
SulcalDepth-entropy	0.00346167	0.310751
FreeSurfer	0.075644	0.303376

Fig. 2. Mean sample variances of sulcal depth and cortical thickness measurements across the cortical surface, given different correspondence maps.

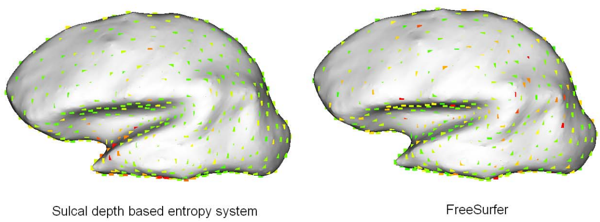

Fig. 3. Comparison of the distribution of variance across the cortical surfaces for our method (left) and FreeSurfer (right). The coloring of the particles is linearly growing from values of (green=0) to (red=1). It can be seen that the particle system has a localized high variance near the temporal lobe (conceivably due to reconstruction noise), but has low variance elsewhere; FreeSurfer has relatively high variance across the entire cortical surface.

and standard entropy-based particle system (location only) results, as well as FreeSurfer results. Mean sample variances for both the cortical thickness and sulcal depth measurements are summarized in Fig. 2.

For sulcal depth measurements, our method reduces variance almost 75-fold over initial data, and almost 25-fold over FreeSurfer results. For cortical thickness, our method has considerable improvement over initial data, and has a slightly higher average variance than FreeSurfer does. An inspection of the distribution of this mean variance over the surface, as shown in Fig. 3, reveals that FreeSurfer has a higher variance across the entire surface, whereas our method has only a localized high variance around the temporal lobe (which is, as noted above, not perfectly reconstructed due to input image noise), and performs much better in other areas of the cortical surface.

It should be noted that both our method and FreeSurfer use the sulcal depth information as part of the correspondence optimization process; therefore, the sulcal depth evaluation is biased. Also note that the higher degree of remaining variability in cortical thickness can be largely attributed to inter-subject variability, as cortical thickness patterns tends to vary largely among individuals.

5. CONCLUSIONS

Our method provides a novel way of computing correspondence across datasets of human cortical surfaces using functions of spatial locations in an entropy minimizing particle framework. We show that our method provides correspondence maps that result in tighter distributions of sulcal depth and cortical thickness compared to other commonly used methods such as FreeSurfer. The cortical thickness measurements show some locally high variances near the temporal lobe, where the FreeSurfer reconstruction was less than perfect.

This framework provides exciting new directions of research. Integration of fiber connectivity and vasculature information, extracted from DTI and MRA scans respectively, will lead to a better understanding of cortical correspondence, and will provide further flexibility in the definition of correspondence, depending on the specific target application. Our next steps involve experimenting the integration of probabilistic connectivity maps that reflect the fiber structures between the cortical surface and various regions of interest, such as the corpus callosum and the internal capsule.

6. REFERENCES

[1] J. Cates, T. Fletcher, and R. Whitaker, "Entropy-based particle systems for shape correspondence," *Mathematical Foundations of Computational Anatomy Workshop, MICCAI 2006*, pp. 90–99, Oct. 2006.

[2] A. C. Kotcheff and C. J. Taylor, "Automatic construction of eigenshape models by direct optimization," *Medical Image Analysis*, vol. 2, no. 4, pp. 303–314, December 1998.

[3] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modelling," *IEEE trans. Medical Imaging*, vol. 21, no. 5, pp. 525–537, 2002.

[4] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "3D statistical shape models using direct optimisation of description length," in *Proc. European Conference on Computer Vision, Part III*. 2002, pp. 3–20, Springer.

[5] C. Brechbühler, G. Gerig, and O. Kübler, "Parametrization of closed surfaces for 3-D shape description," *Computer Vision and Image Understanding*, vol. 61, pp. 154–170, March 1995.

[6] T. Heimann, I. Oguz, I. Wolf, M. Styner, and H. P. Meinzer, "Implementing the automatic generation of 3d statistical shape models with itk," *ISC/NA-MIC Workshop on Open Science at MICCAI 2006*, 2006.

[7] M. Styner, J. Lieberman, and G. Gerig, "Boundary and medial shape analysis of the hippocampus in schizophrenia," .

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley-Interscience, August 1991.

[9] J. Cates, T. Fletcher, M. Styner, M. Shenton, and R. Whitaker, "Shape modeling and analysis with entropy-based particle systems," in *IPMI*, 2007, pp. 333–345.

[10] F. Segonne, A. M. Dale, E. Busa, M. Glessner, DU Salvolini, H. K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in mri," *NeuroImage*, vol. 22, pp. 1060–1075, July 2004.

[11] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis i: Segmentation and surface reconstruction," *NeuroImage*, vol. 9, pp. 179–194, 1999.

[12] B. Fischl, M. I. Sereno, and A. M. Dale, "Cortical surface-based analysis ii: Inflation, flattening, and a surface-based coordinate system," *NeuroImage*, vol. 9, pp. 195–207, 1999.

[13] B. Fischl, A. Liu, and A. M. Dale, "Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 1, pp. 70–80, Jan 2001.

[14] B. Fischl, M. I. Sereno, R. B. H. Tootell, and A. M. Dale, "High-resolution inter-subject averaging and a coordinate system for the cortical surface," *Human Brain Mapping*, vol. 8, pp. 272–284, 1999.

[15] B. Fischl and A. M. Dale, "Measuring the thickness of the human cerebral cortex from magnetic resonance images," in *Proceedings of the National Academy of Sciences*, 2000, vol. 97, pp. 11044–11049.