Automated Diagnosis of Autism Using Fourier Series Expansion of Corpus Callosum Boundary

Shubing Wang[1,5], Moo K. Chung[1,2,5], Kim M. Dalton[5], Andrew L. Alexander[3,5] and Richard J. Davidson[4,5]

[1] Department of Statistics, [2] Department of Biostatistics and Medical Informatics, [3] Department of Medical Physics,
[4] Department of Psychology and Psychiatry, [5] Waisman Laboratory for Brain Imaging and Behavior,
University of Wisconsin at Madison

Abstract

We explored the possibility of developing an automatic diagnostic tool for detecting autism based on MRI measurements. Since the two previous structural imaging studies [1] [2] strongly suggested there were significant abnormality in the corpus callosum (CC) region, the methodology is concentrated in this area. For this purpose, we have developed a new framework for representing and classifying the CC boundary as a Fourier descriptor [3] [6]. The Fourier coefficients can be viewed as a multivariate measurement that characterizes the CC boundary, and later feed into a classification algorithm.

Segmentation

Three Tesla T1-weighted MR scans were acquired for 15 high functioning autistic and 12 control right-handed males. The CC boundaries were extracted using the gradient vector flow (GVF) snakes [6] on the mid-saggital sections of MRI. The GVF snakes have advantages over the traditional snakes with their larger capture range and ability to move into the concavity. (Figure 1).

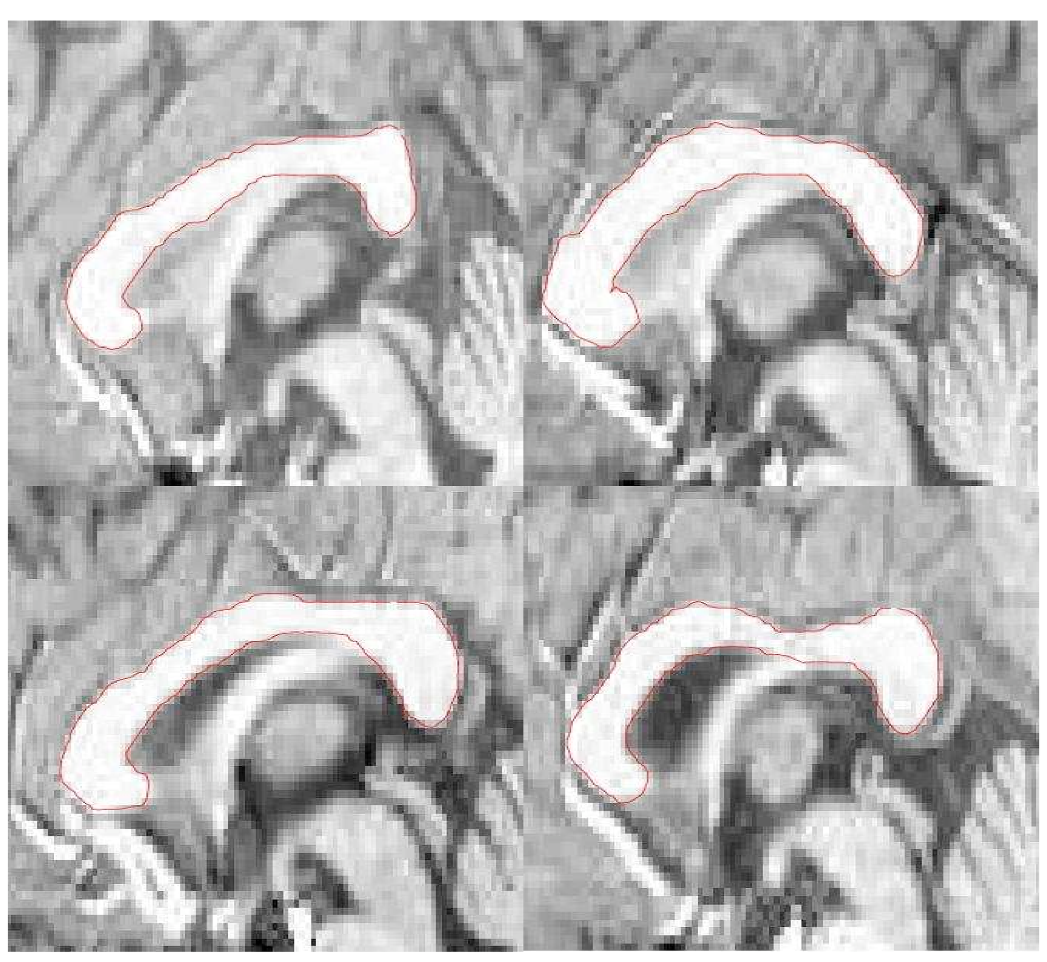

Figure 1. Results of CC boundaries by GVF snakes.

Parametrization

The extracted CC boundaries were parameterized by arc-length. Classical curvature calculation uses the values of first and the second derivatives, and therefore need a parametrization of the given curves. A new algorithm for computing the curvature, which is invariant under parametrization, is developed. An obtained CC boundary is a discrete curve $\{P_i\}_{i=1}^n$. Let $A(P_{i-1}, P_i, P_{i+1})$ be the area of triangle with vertices P_{i-1}, P_i, P_{i+1} and "sign" is 1 if the triangle is inside the CC boundary and -1 otherwise. Curvature at P_i can be calculated using (as shown in Figure 2)

$$\hat{k}_i = \text{sign} \cdot \frac{4A(P_{i-1}, P_i, P_{i+1})}{\|P_{i-1} - p_i\| \cdot \|P_{i+1} - P_i\| \cdot \|P_{i+1} - P_{i-1}\|} \qquad (1)$$

a simple closed (no-cross) curve with absolutely continuous second derivatives, let

$$h = \max_{i=1,2,\cdots,n-1} \|P_{i+1} - P_i\|. \qquad (2)$$

If the CC boundary is a simple noncrossing closed curve with absolutely continuous second derivatives, then it has an absolutely continuous curvature function k. We can show that

$$\|\hat{k} - k\|_\infty = O(h \log(h)), \quad \text{as } h \to 0.$$

where $\|\cdot\|_\infty$ is the L^∞-norm.

A simple arc-length parametrization $\{\boldsymbol{r}(s_i) = P_i\}, i = 1, 2, \cdots, n$ of an obtained CC boundary $P_{i=1}^n$ is given by

$$s_1 = 0$$
$$s_i = \sum_{l=1}^{i-1} \|P_{l+1} - P_l\|, \quad i = 2, 3, \cdots, n.$$

This parametrization can be a good approximation if the curvature is small (Figure 2, bottom right), while it also can be a very poor approximation if the curvature is large (Figure 2, top right).

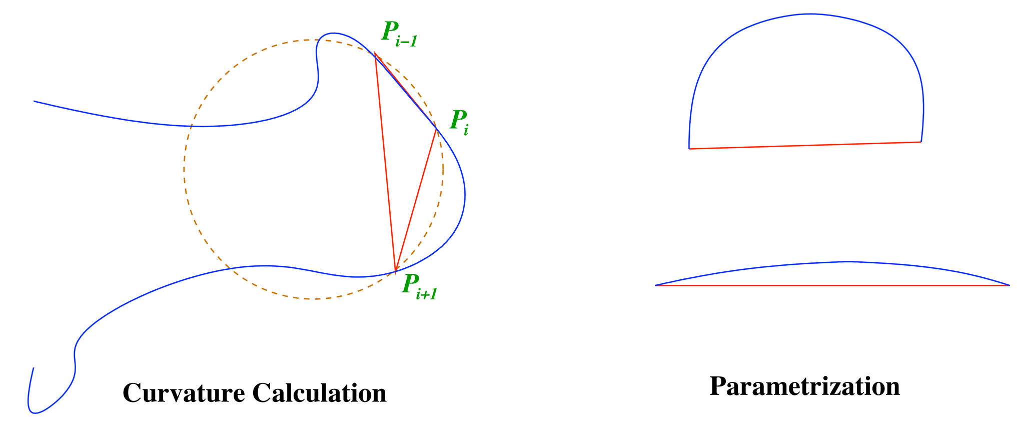

Figure 2. Left, the demonstration of how to calculate curvature; Right comparison of distance of two points with the arc-length between two points: top, when curvature is large; bottom, when curvature is small.

We designed an algorithm of arc-length parametrization utilizing curvature information, which gives a more accurate approximation. Let $\boldsymbol{r}(s)$ be the parametrization of a CC boundary, $\boldsymbol{T}(s)$ be the unite tangent vector and $\boldsymbol{N}(s)$ be the unit normal vector. Based on the Frenet formulae

$$\begin{cases} \dot{\boldsymbol{r}}(s) = \boldsymbol{T}(s) \\ \dot{\boldsymbol{T}}(s) = k(s)\boldsymbol{N}(s) \end{cases}$$

we designed the following parametrization method

$$s_1 = 0$$
$$s_i = s_{i-1} + \|P_i - P_{i-1}\| + \lambda\sqrt{k_i \|\boldsymbol{P}_{i-1} + \boldsymbol{P}_{i+1} - 2\boldsymbol{P}_i\|}.$$

where λ is a parameter that controls the correction factor, and curvature k_i is required to be independent of parametrization. Using (1) is an advantage over the classical method. With the higher order correction factor using curvature information, our parametrization has order of convergence $o(h^2)$, while the simple parametrization method only has order of convergence $o(h)$, where h is defined in (2). Figure 3 shows that the simple parametrization is under-estimated, but our method gives a better parametrization (closer to the ground truth).

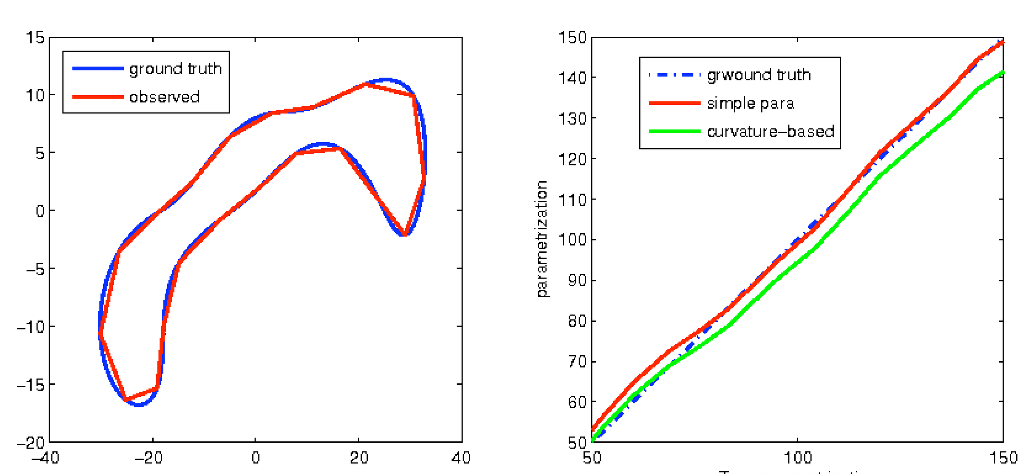

Figure 3. Left, simulated CC boundaries; Right, the comparison of two parametrization results versus true parametrization.

Curvature-based registration

Curvature function $\{k(s_i)\}_{i=1}^n$ of closed curve $\{\boldsymbol{r}(s_i) = (x(s_i), y(s_i))^\tau\}_{i=1}^n$ is invariant under a rigid body motion. The corresponding closed curve can be reconstructed from the curvature function by

$$x(s) = x(s_1) + \int_{s_1}^{s} \cos(\theta(s))ds,$$
$$y(s) = y(s_1) + \int_{s_1}^{s} \sin(\theta(s))ds,$$

where $\theta(s) = \int_{s_1}^{s} k(s)ds$.

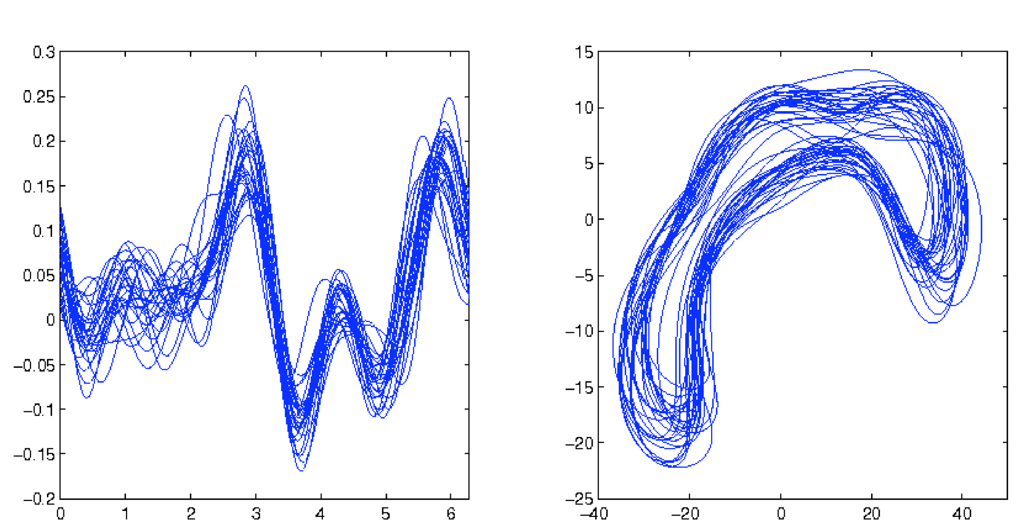

Figure 4. Left: global shift registration results of curvature functions; right: registered closed curves recovered by registered curvature functions.

The global shift registration technique [5] is used to align 27 CC curves. Aligned curves $\{k_i^*(s) = k_i(s + \delta_i)\}_{i=1}^{27}$ minimize

$$\text{REGSSE} = \sum_{i=1}^{27} \int_0^{2\pi} [k_i(s + \delta_i) - \hat{\mu}(s)]^2 ds$$
$$= \sum_{i=1}^{27} \int_0^{2\pi} [k_i^*(s) - \hat{\mu}(s)]^2 ds$$

where $\hat{\mu}(s)$ is the mean curve of $\{k_i^*(s)\}_{i=1}^{27}$. The optimal aligned curves minimize the summation of sum of square errors (SSE) of aligned curves from their mean curve.

Classification

As a data reduction technique, the CC curves were reparametrized by the finite Fourier series on the coordinate functions. The Fourier coefficients were used as predictors in classifying the CC curves into two groups: normal controls and autism. Decision-tree-based classification techniques [3] were applied to determine if it is possible to differentiate autism purely based on the shape of CC curves. Decision-tree-based methods have been widely used in statistical literatures as a nonparametric method and there is no explicit statistical assumptions about predictors and response variables. When the classification boundary is expected to be nonlinear (Figure 5 left), the regression tree methods should perform better than the linear discriminant analysis (Figure 5, right).

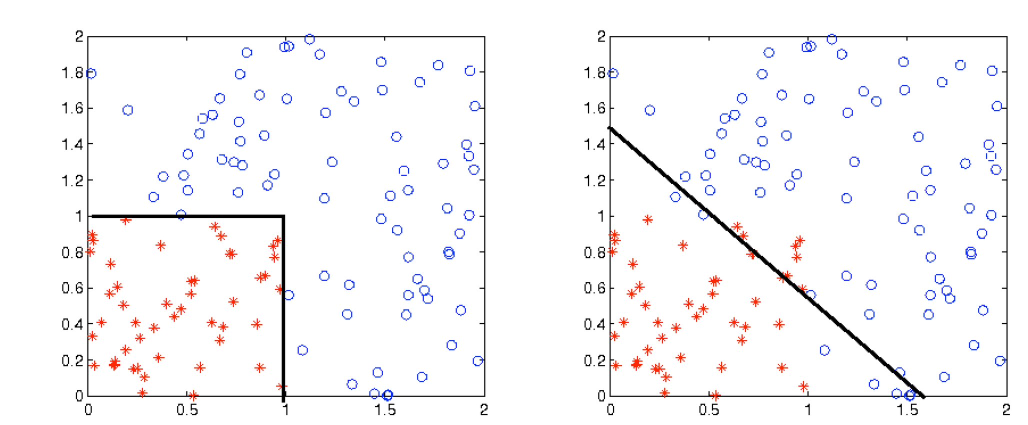

Figure 5. Left: classification using regression trees. Right: linear discriminant analysis classification.

Results

Regression tree methods [4]: CRUISE, GUIDE are used, Linear Discriminant Analysis (LDA) is also used for comparison.

Methods	LDA	CRUISE	GUIDE
Misclassification rate	0.25	0.22	0.15

Table 1. The classification results of various methods.

With a small sample size of 27 subjects, we still managed to achieve an impressive 15% misclassification rate (85% correct diagnostic rate) consistent with the result of two previous structural imaging studies done on the CC [1] [2]. With the additional social and behavioral measurements, the correct diagnostic rate can be improved.

References

[1] Aleander et al. (2007). Diffusion tensor imaging of the corpus callosum in autism. *NeuroImage*, 34:61-73.

[2] Chung et al. (2003). Less white matter concentration in autism: 2D voxel-based morphometry. *NeuroImage* 23:242-251.

[3] Golland, P., Frimson, W.E.L. and Kikinis, R. (1999). Statistical Shape Analysis Using Fixed Topology Skeletons: Corpus Callosum Study. *Proc. IPMI'99: 16th International Conference on Information Processing and Medical Imaging*, LNCS 1613, pp. 382-387, Springer-Verlag.

[4] Loh, W.Y. (2002). Regression trees with unbiased variable selection and interaction detection, *Statistica Sinica*, 12:361-386.

[5] Ramsay, J.O. and Silverman, B.W. (1997). *Functional Data Analysis*, 2nd Edition, Springer.

[6] Xu, C. and Prince, L. (1998). Snakes, shapes and gradient vector flow, *IEEE Trans. on Image Proc.*, 7:359-369.

[7] Zahn C.T. and Roskies, R.Z. (1972). Fourier descriptors for plane close curves, *IEEE Trans. Computers*, Vol C-21, pp. 269-281.