

Pattern classification using principal components of cortical thickness and its discriminative pattern in schizophrenia [☆]

Uicheul Yoon,^a Jong-Min Lee,^{a,*} Kiho Im,^a Yong-Wook Shin,^b Baek Hwan Cho,^a
In Young Kim,^a Jun Soo Kwon,^b and Sun I. Kim^a

^aDepartment of Biomedical Engineering, Hanyang University, Sungdong PO Box 55, Seoul 133-605, Korea

^bDepartment of Psychiatry, Seoul National University College of Medicine, Seoul, Korea

Received 23 March 2006; revised 2 October 2006; accepted 7 November 2006
Available online 26 December 2006

We proposed pattern classification based on principal components of cortical thickness between schizophrenic patients and healthy controls, which was trained using a leave-one-out cross-validation. The cortical thickness was measured by calculating the Euclidean distance between linked vertices on the inner and outer cortical surfaces. Principal component analysis was applied to each lobe for practical computational issues and stability of principal components. And, discriminative patterns derived at every vertex in the original feature space with respect to support vector machine were analyzed with definitive findings of brain abnormalities in schizophrenia for establishing practical confidence. It was simulated with 50 randomly selected validation set for the generalization and the average accuracy of classification was reported. This study showed that some principal components might be more useful than others for classification, but not necessarily matching the ordering of the variance amounts they explained. In particular, 40–70 principal components rearranged by a simple two-sample *t*-test which ranked the effectiveness of features were used for the best mean accuracy of simulated classification (frontal: (left(%)|right(%))=91.07|88.80, parietal: 91.40|91.53, temporal: 93.60|91.47, occipital: 88.80|91.60). And, discriminative power appeared more spatially diffused bilaterally in the several regions, especially precentral, postcentral, superior frontal and temporal, cingulate and parahippocampal gyri. Since our results of discriminative patterns derived from classifier were consistent with a previous morphological analysis of schizophrenia, it can be said that the cortical thickness is a reliable feature for pattern classification and the potential benefits of such diagnostic tools are enhanced by our finding.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Magnetic resonance imaging; Cortical thickness; Principal component analysis; Support vector machine; Leave-one-out cross-validation

Introduction

Recent advances in magnetic resonance image acquisition and processing have allowed for the morphometric analysis of the cerebral cortex at a macroscopic level and have allowed an investigation of normal and abnormal changes. Much structural magnetic resonance imaging (MRI) of the brain in schizophrenia has indicated subtle cortical abnormalities compared with healthy controls. Specifically, gray matter (GM) deficits in superior and medial temporal cortices have been widely reported, but there is moderate evidence of focal GM volume reductions in frontal, parietal, and occipital neocortices, and subcortical abnormalities (Shenton et al., 2001). Although voxel-based morphometry (VBM), which involves a voxel-wise comparison of local GM concentration, is the most widely used approach, it might in part reflect differences in the surrounding tissue and might be influenced by the sulcal widening in schizophrenia (Ashburner and Friston, 2000; Narr et al., 2005a; Shenton et al., 2001). Although GM concentration reflects the proportion of GM within cortical mantle with respect to other tissue types, and cortical thickness represents the distance across the cortex according to some geometric definition, it was reported that these measures are highly correlated, but none of several potential defects in GM concentration were associated with the cortical thickness (Narr et al., 2005a). Therefore, it can be said that measuring cortical thickness provides a closer approximation to the underlying anatomical reality and a direct quantitative index of cortical morphology.

Several postmortem studies have assessed cortical thickness in schizophrenia, but they are limited by labor-intensive procedures, making it impractical to measure cellular density and thickness in all cortical regions (Selemon, 2004; Selemon et al., 1995). However, *in vivo* data with the latest computational neuroimage analysis methods may allow differences in cortical thickness to be estimated from the nodes of a 3D polygonal mesh rather than from a 3D voxel grid. The surface-based approach has the following additional advantages over image- or voxel-based approaches. First, it can be applied in more general situations where a surface is not embedded in an image, but is defined in another way such as segmented boundaries or triangulations. Second, unless the appearance

[☆] Pattern classification using cortical thickness.

* Corresponding author. Fax: +82 2 2296 5943.

E-mail address: ljm@hanyang.ac.kr (J.-M. Lee).

URL: <http://cna.hanyang.ac.kr> (J.-M. Lee).

Available online on ScienceDirect (www.sciencedirect.com).

rance inside the object is also the focus of interest, it may be more appropriate for shape analysis, as the boundary or surface of a volumetric object actually defines the shape. Finally, some noise generated from resampling in the voxel-based analysis can be avoided. Few prior surface-based approaches have examined cortical thickness in schizophrenia (Kuperberg et al., 2003; Narr et al., 2005a,b; White et al., 2003; Wiegand et al., 2004). Kuperberg et al. (2003) assessed thinning across the entire cortex and showed widespread significant thinning that particularly affected the prefrontal and temporal cortices in chronic schizophrenia. White et al. (2003) reported significant cortical thinning in cortex underlying the sulci in frontal, temporal and parietal regions and beneath the gyri in the temporal lobe in patients with childhood and adolescent onset schizophrenia. Another study examining cortical thickness averaged across the entire prefrontal cortex failed to detect significant cortical thinning in first-episode schizophrenia (Wiegand et al., 2004). Narr et al. (2005a) revealed significant regional GM thinning in the frontal, temporal and parietal heteromodal association cortices bilaterally in first-episode schizophrenia. Therefore, it could be expected that these findings related to differences of cortical thickness as a quantitative index of cortical morphology contribute affirmatively toward a categorization between schizophrenia and healthy control.

Because schizophrenia is a large-scale disorder of neurocognitive networks rather than confined to specific regions, and structural changes are present in multiple brain regions, it can be anatomically characterized by abnormality at a supra-regional level of brain organization (Burns et al., 2003; Lawrie and Abukmeil, 1998; Shenton et al., 2001; Wright et al., 1999b). Based on previous findings, Davatzikos et al. (2005) performed whole-brain analysis of structural differences between schizophrenic patients and healthy controls, and applied a high-dimensional nonlinear pattern classification technique to quantify the degree of separation between patients and controls. They were able to classify new individuals as schizophrenic or healthy with 81% accuracy and suggested the potential utility of MRI as a diagnostic aid. Principal component analysis (PCA) is a multivariate method identifying correlation among a set of measurements or variables so that it obtains a set of basis vectors whose linear combination can optimally represent the measured data. And it may also be used to obtain a low-dimensional representation of the measurements themselves. Even if the results may be unstable where the number of subjects is much smaller than the number of variables, dimensionality reduction of PCA is very effective in classification because a higher number of features will easily lead the classifier into the problem of overfitting. Initially, Olson and Miller (1958) proposed that anatomical structures could be recomposed into supra-regional systems by PCA of the covariance or correlations between regional elements (Wright et al., 1999b). Since then, regional elements of supra-regional systems defined by PCA have often been found to share developmental influences or to have a common function (Cheverud, 1982). Narr et al. (2005a) used PCA to reduce cortical thickness measures into principal components, but they examined global effect of cortical GM. Although several shape classification studies have been conducted for discovering hippocampal shape abnormality in schizophrenia, to our knowledge this is the first classification study based on principal components of cortical thickness measured at the spatially homologous cortical surface locations in each individual.

The aim of classification is to instruct the classifier using a training set or set of labeled examples representing different

classes, and then use the classifier to predict the class of any new example. This constitutes the final goal of the learning stage in many application domains, including character recognition and text classification. In medical image analysis, however, it has been much more useful in understanding the nature of the differences captured by the classifier than in using it for labeling new examples. These differences, expressed in terms of the original images or shapes, can provide an insight into the anatomical implications of shape differences detected by the learning algorithm. Furthermore, it could be argued that studying the structure of the data captured by the classifier is important in any application, because it puts an emphasis on the nature of the differences between the classes and can potentially help improve the technique. Golland et al. (2005) introduced the notion of the *discriminative direction* at every point in the feature space with respect to a given classifier, which corresponds to the maximum changes in the classifier's response while minimizing irrelevant changes in the input. It allows to characterize shape differences between the two classes captured by the classifier and to express them as deformations of the original shape. Shen et al. (2004) proposed discriminative patterns that shared a similar idea of discriminative direction: for a linear classifier, the deformation representing class differences could be visualized using the normal to the separating hyperplane. In this study, we proposed pattern classification based on principal components of cortical thickness between schizophrenic patients and healthy controls, and validated its accuracy using a leave-one-out cross-validation (LOOCV) method. Discriminative patterns derived at every vertex in the original feature space with respect to a given classifier were analyzed with definitive findings of brain abnormalities in schizophrenia for establishing practical confidence. The purpose of this study is identification of representative regions contributing to the classification through a discriminative pattern. Although it is less conservative than statistical *t*-test, our method would be useful to grasp the trend of difference between healthy control and schizophrenia in cortical thickness.

Materials and methods

Subjects

A group of right-handed schizophrenic patients was recruited from the inpatient unit and the outpatient clinic at Seoul National University Hospital, Seoul, Korea. Fifty-three patients (32 men, 21 women) were interviewed using a structured clinical interview based on DSM-IV (SCID-IV) and met those criteria for schizophrenia. The Structured Clinical Interview (SCID) was first devised with the DSM-III-R (APA, 1987), and it was modified in 1997 to conform to the new DSM-IV system of diagnosis (First et al., 1997). The SCID is well designed to diagnose schizophrenia, which also separates its main diagnosis from other psychiatric disorders since it has a form of a branching-style interview covering most major psychiatric symptoms. It is noteworthy here that since the introduction of SCID, almost all studies of schizophrenia have been using it for the diagnosis of schizophrenia. Exclusion criteria for patients were any lifetime history of neurological or significant medical illnesses, and any past history of substance abuse. Each patient's symptoms were rated on the positive and negative syndrome scale (PANSS, Kay et al., 1987). The normal control group was recruited from internet advertise-

ments, and consisted of 52 healthy subjects matched with the patient group for age, sex, handedness, and socioeconomic status (SES). Exclusion criteria for controls were any current or lifetime history of a DSM-IV axis I disorder. The demographic characteristics of patients and controls are summarized in Table 1. The mean PANSS total score of the patients was 72.26 ± 15.38 . This study was carried out under the guidelines for the use of human subjects established by our institutional review board. All subjects gave written informed consent for the procedures before their participation in the study.

MR image acquisition and cortical surface extraction

MR images were acquired using a 1.5-T GE SIGNA scanner (GE Medical Systems, Milwaukee, WI) with a 3D-SPGR T1-weighted spoiled gradient echo pulse sequence with the following parameters: 1.5 mm sagittal slices; echo time 5.5 ms; repetition time 14.4 ms; number of excitations 1; rotation angle 20° ; field of view 21×21 cm; matrix 256×256 voxels. Images were resampled to be isocubic and realigned so that the anterior–posterior axis of the brain was aligned parallel to the intercommissural line and the other two axes were aligned along the interhemispheric fissure. The datasets were then filtered using an affine anisotropic diffusion filtering to improve the signal-to-noise ratio, which is a scale space, adaptive and the most widely used nonlinear technique for noise reduction. These procedures were processed using ANALYZE 4.0 software (Mayo Foundation, USA).

The following steps were applied to these data, which have been described in detail elsewhere (Collins et al., 1994; Kim et al., 2005; MacDonald et al., 2000; Sled et al., 1998; Zijdenbos et al., 2002). A fully 3D technique for inhomogeneity correction removed a serious obstacle for automated segmentation of MRI, which involved slowly varying the change in signal intensity over the image caused by magnetic field inhomogeneity (Sled et al., 1998). To account for inter-individual differences in absolute brain size, each brain was separately transformed into a standardized stereotaxic space and resampled on a 1 mm^3 voxel grid. This was done using automatic registration software that used a 3D cross correlation approach to match the single MRI volume with the intensity average of 305 MRI brain volumes previously aligned in standardized stereotaxic space (Collins et al., 1994). An artificial neural network classifier was applied to identify GM, white matter (WM) and cerebrospinal fluid (CSF) (Zijdenbos et al., 2002).

Cortical surfaces were automatically extracted from each MR volume using the Constrained Laplacian-based Automated Segmentation with Proximities (CLASP) algorithm, which reconstructed the inner cortical surface by deforming a spherical mesh onto the WM/GM boundary (Kim et al., 2005; MacDonald et al., 2000). Hemispheric cortical surface models were constructed; these consisted of 81,920 polygons forming high-resolution meshes of discrete triangular elements.

Measurement of cortical thickness and principal component analysis

Because the cortical surface models were extracted from MR volumes transformed into stereotaxic space, the inverse transformation matrix was applied to cortical surfaces for measuring cortical thickness in native space. Inner and outer surfaces had the same vertex number, and the correspondence between surfaces in each vertex was defined. Thus, cortical thickness was easily measured using the T_{link} method of calculating the Euclidean distance between linked vertices on the WM/GM boundary surface and the GM/CSF intersection surface (Kabani et al., 2001; Lerch and Evans, 2005; MacDonald et al., 2000). Although there are several algorithms for measuring cortical thickness, evaluation of their precision suggested that the T_{link} method is the simplest and most precise method (Lerch and Evans, 2005). To apply PCA to the cortical thickness of whole subjects, the thickness information was spatially normalized using surface-based registration. Since the cortical surface of each subject starts from a spherical polygon model in the CLASP algorithm, the vertices are easily transformed to the spherical model and nonlinearly registered to an average template on the sphere. This 2D surface registration makes use of a highly flexible deformation of a template cortex to an individual (Robbins, 2003). It provides a geodesic distance transformation to match crowns of gyri between subjects. This algorithm was tuned for chosen parameter values, improving the resulting registrations. Sulcal variability was reduced in all areas of the cortex using optimal parameter values, which was proven by the method of entropy measure (Robbins et al., 2004). Finally, thickness information on each vertex of whole subject was transformed to a template so that the correspondence between subjects at each vertex of the cortical surface model should be assured. This algorithm has been used very well for previous studies which compared the thickness of the surface model between groups (Im et al., 2006; Lerch et al., 2006). Then, diffusion smoothing, which generalizes Gaussian kernel smoothing, with 20 mm FWHM (full width half maximum) was used to increase the signal to noise ratio and detect population changes very well. It was chosen as the kernel size to maximize statistical power while still minimizing false positives (Chung et al., 2002; Lerch et al., 2005).

The 40,962 vertices of the hemispheric cortical surface model generated the same number of variables so that implementing PCA over that immense number of variables was impractical because of computer memory restrictions. In addition, the number of variables was much larger than the number of subjects, so the results of PCA might lack stability. The brain can be divided into four main areas: the frontal, parietal, temporal, and occipital lobes. In this study, the average cortical surface was generated from manually parcellated cortical surfaces of 34 subjects, consisting of 17 healthy control and 17 schizophrenic patients using *SUMA* software (available at <http://afni.nimh.nih.gov>) and used as a template for the lobar definition of each subject (Cox and Hyde, 1997). The definition of

Table 1
Demographic characteristics of the subjects

Characteristic	Schizophrenia (N=53)		Healthy control (N=52)	
	Mean	SD ^a	Mean	SD
Age (years)	27.80	6.00	26.44	5.35
Sex (male/female)	32/21		32/20	
Subjects educational level (years)	14.37 *	2.65	15.67 *	2.01
Socioeconomic status	3.04 *	0.88	2.65 *	0.59
Age of onset (years)	23.96	6.18	N/A	
Duration of illness (years)	4.91	4.09		
PANSS ^b total scores	72.40	15.70		

^a Standard deviation.

^b Positive and negative syndrome scale.

* $p < 0.01$.

the boundary between lobes (Fig. 1) was described in detail in our previous publication (Yoon et al., 2005).

The input data for PCA become a different number of elements according to lobes. Given a group of N subjects, the mean data \bar{X}_j of lobe j can be calculated using

$$\bar{X}_j = \frac{1}{N} \sum_{i=1}^N X_{ij},$$

where X_{ij} is the lobe j of the i th subject.

Subsequently, PCA by lobe was applied to reduce dimensionality and make classification feasible. The subscription for lobe will be left out to simplify the equation. This involves eigenanalysis of the covariance matrix Σ of the data as follows:

$$\Sigma = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})^T,$$

$$\Sigma E = \Lambda E,$$

where the columns of E hold eigenvectors, and the diagonal matrix Λ holds eigenvalues of Σ . The eigenvectors can be ordered according to respective eigenvalues, which describe the amount of variance dictated by each eigenvector. On the other hand, one particular lobe of a subject can then be written

$$b_i = E^T(X_i - \bar{X})$$

where b_i is a vector containing principal components representing the difference between X_i and \bar{X} in the eigenvector space, which can be thought of a new and more compact representation of the data. Given a dataset of N subjects, the first $N-1$ principal components are enough to capture all the data variance. PCA can also be viewed as a kind of feature extraction from the original data. In this study, these principal components were used as a feature vector of the following classifier.

Support vector machine and feature selection methods

The support vector machine (SVM) belongs to a learning system based on recent advances in statistical learning theory (Vapnik, 1998). We applied linear SVM using LIBSVM, which is a library for support vector machines (Fan et al., 2005). Although we have briefly described how to train linear classifiers with SVM, an extensive introduction and a formal discussion of the theoretical

foundations of SVM can be found in the literature (Burges, 1998; Vapnik, 1998).

Given a training set of principal components and label pairs (b_i, y_i) , $i=1,2, \dots, N$ where $y_i \in \{1,-1\}$, the SVM requires the solution of the following optimization problem:

$$\min_{\omega, b, \xi} \frac{1}{2} \omega^T \omega + C \sum_{i=1}^N \xi_i$$

subject to

$$y_i(\omega^T \phi(b_i) + \beta) \geq 1 - \xi_i, \quad \xi_i \geq 0.$$

Training vectors b_i are mapped into a higher (maybe infinite) dimensional space by the function ϕ . Therefore, ω is normal to the hyperplane and $|\beta|/|\omega|$ is the perpendicular distance from the hyperplane to the origin. In addition, ξ_i is a non-negative slack variable for non-separable cases and C is a parameter for adjusting the cost of the constraint violation, i.e., the compromise between maximizing the margin which can be defined as the sum of the distances from the hyperplane to the closest positive and negative exemplars, and minimizing the number of errors. Then SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. A linearly separable SVM aims to find the separating hyperplane with the largest margin.

Classification is performed on feature vectors after PCA and LOOCV methodology is used for finding the best parameters of classifier (Duda et al., 2000). As the name suggests, LOOCV involves using a single observation from the original sample as test data, and the remaining observations as training data. This is reiterated such that each observation in the sample is used as the test data. In this study, in order to keep away from any other bias, validation set for classifier evaluation was separated from training and test set for classifier learning. The validation set was randomly selected from the whole subject and consisted of 15 healthy control and 15 schizophrenic patients. The parameter of linear SVM classifier based on each feature vector was determined in case of the best leave-one-out accuracy from the remaining subjects as a training and test set, and then the performance of this classifier was evaluated using the validation set. For the generalization, it was simulated with 50 randomly selected validation set and the average accuracy of classification was reported. Principal components, whose standard ordering is by the variance amounts they describe, are generally used as features in the classification, and more features can usually improve the performance of the classifier. However, each additional feature supplements a parameter to the estimated model, and incorrect estimations from the less informative features can degrade performance. This trend of decreasing accuracy gains followed by actual losses of accuracy from additional features is known as the Hughes phenomenon (Hughes, 1968). Therefore, it is helpful to select a subset of the most useful features. Shen et al. (2004) reported that some principal components are more useful than others for classification, but not necessarily for matching the ordering of the variance amounts they describe. A simple two-sample t -test was used on each feature to rank the effectiveness of features. A lower P value implies a statistically stronger group difference and therefore corresponds to a more significant feature. Two types of feature selection schemes were investigated in this study, which were orderings of principal components by variance explained and P value associated with the t -test.

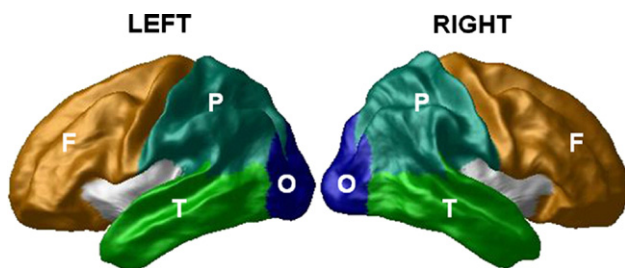


Fig. 1. The color-coded definition of the boundary between lobes was mapped onto the average cortical surface. It was used as a template for the lobar definition of each subject.

Discriminative patterns for classification

An approach for visualizing discriminative patterns identified by a linear classifier was provided for a comprehensible description

of the group difference. We only state without proof the main facts of this method that are necessary for derivation of our results, so that the reader is referred to the original publications for more details (Golland et al., 2005; Shen et al., 2004).

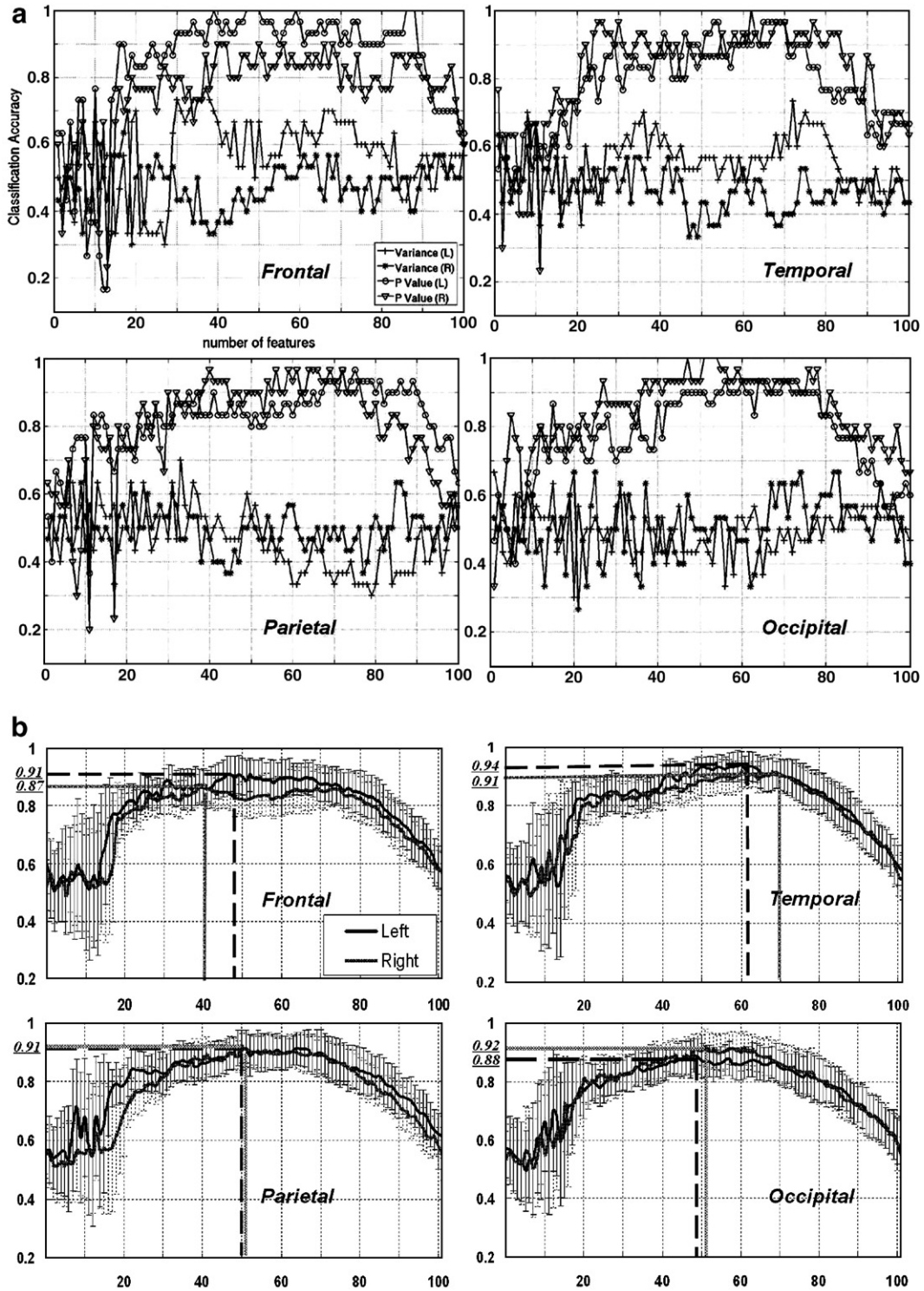


Fig. 2. Classification results on each lobe (a) different feature selection schemes: orderings of principal components by variance explained and *P* value associated with two-sample *t*-test; (b) mean and standard deviation for 50 randomly selected validation set. The *x*-axis is the number of features used for the classification and the *y*-axis is the accuracy of classification.

Applying PCA and SVM as detailed above to a shape set, a discriminative value V_i for each subject X_i could be calculated as follows:

$$V_i = (X_i - \bar{X})^T * E * \omega = (X_i - \bar{X})^T * P_D,$$

where P_D is a column vector that weights the contribution of each difference in each subject to a discriminative value. Given a vertex k of the cortical surface model, we use $P_D(k)$ to denote the corresponding weights for difference of thickness. As there is no information of direction on the shape used in this study, only an absolute value of $P_D(k)$ is valid. A large magnitude of $P_D(k)$ indicates that vertex k has discriminative power, as even small local differences at this vertex will have a noticeable effect on the overall classification. In a way of showing statistical group difference implied by the classifier model, the color-coded value of P_D in each vertex was mapped onto a mean surface to identify significant discriminative regions. In this study, discriminative patterns were obtained by the following procedure:

- (1) PCA was applied to cortical thickness of each lobe;
- (2) principal components were ordered by P value associated with the t -test and used as features of SVM classifier;
- (3) SVM using the minimum number of features needed was applied to achieve a perfect discrimination between classes;
- (4) discriminative pattern was calculated using eigenvectors depending on selected principal components and the corresponding normal to the separating hyperplane;
- (5) discriminative pattern for each case of the simulation was averaged vertex-by-vertex in the original space and mapped onto a cortical surface;
- (6) anatomical labels of representative regions were acquired using the *Talairach Daemon Client* (version 2.0, available at <http://ric.uthscsa.edu/projects/talairachdaemon.html>).

For better comprehension of discriminative patterns, spatially normalized cortical thickness of patients and controls were compared statistically using vertex-wise t -tests. Correction for multiple comparisons was needed to control the false-positive rate. The false discovery rate (FDR) controlling procedure for multiple comparisons was reported to be effective for the analysis of neuroimaging data (Genovese et al., 2002). We performed FDR correction for multiple comparisons at $P < 0.05$.

Results

Classification results using different feature selection schemes

Classification can be performed using just the first few principal components, which account for significant amount of data variance, based on the hypothesis that this information is crucial for classification and the rest noisy. However, judging from the following results, it seems not to be acceptable. Fig. 2(a) shows the classification results on each lobe using different feature selection schemes according to the ordering of principal components: decreasing arrangement by the variance explained (DAV) and increasing one by the P value associated with two-sample t -test (IAP). Table 2 shows the best accuracy of DAV feature selection with the number of features used for each lobe and total variance of those features. In case of the DAV scheme, 63–77% accuracy was achieved when 8–37 principal components

Table 2

Best performance of classifier and the number of features used in the best case

		No. of vertices	DAV ^a		IAP ^b	
			Best accuracy	No. of feature ^c	Best accuracy	No. of feature ^c
Frontal	L	13597	0.77	37 (99.96%)	1.00	39 (98.21%)
	R	13216	0.70	18 (99.70%)	0.90	40 (95.24%)
Temporal	L	6210	0.70	8 (97.27%)	0.97	53 (94.86%)
	R	6242	0.67	10 (98.24%)	1.00	62 (96.86%)
Parietal	L	10525	0.70	33 (99.89%)	0.97	72 (98.23%)
	R	10507	0.63	9 (98.27%)	0.97	40 (93.08%)
Occipital	L	3941	0.63	15 (99.76%)	0.90	46 (95.90%)
	R	4270	0.67	20 (99.83%)	1.00	47 (84.65%)

The numerical value in parentheses represents the variance amounts which the features explain.

^a DAV, decreasing arrangement of features by variance explained.

^b IAP, increasing arrangement of features by P value associated with t -test.

^c It is unnecessary to describe the best accuracy in case of the IAP.

corresponding to 97–99% of the variance were used. It suggested that there were significant differences in cortical thickness between schizophrenic patients and healthy controls. In spite of features covering almost whole variance amounts, however, the accuracy of classification was not sufficient for practical purposes, and additional features degraded the performance of classifier. In contrast, the result of the IAP scheme showed a nearly perfect classification in the best case and always outperformed the DAV for each lobe (Fig. 2(a)). In addition, 39–72 principal components corresponding to 84–98% of the variance were used for the best performance of classifier (Table 2). Note that some components might be more useful than others for classification, but not necessarily match the ordering of the variance amounts they describe. It was noticeable that there was a feature subset supporting nearly perfect classification. Fig. 2(b) shows the average classification accuracy for 50 randomly selected validation set using IAP scheme. On the average by lobe, 40–70 principal components rearranged by IAP scheme were used for the best accuracy of classification (Fig. 2(b), frontal: (left(%)|right(%))=91.07|88.80, parietal: 91.40|91.53, temporal: 93.60|91.47, occipital: 88.80|91.60). And, the result of IAP feature selection always outperformed that of DAV scheme in terms of the best case.

Discriminative patterns

Fig. 3(a) shows the color mapping of discriminative patterns onto a mean surface to identify representative regions contributing to the classification. The regions having more discriminative power are coded in light yellow, and the regions having less are represented by dark red. The representative regions were defined from whole discriminative patterns according to values over a threshold and the size of the cluster. Their anatomical labels and Talairach coordinates obtained by *Talairach Daemon Client* are summarized in Table 3. Discriminative power appeared more spatially diffused bilaterally in the several regions, especially precentral, postcentral, superior frontal and temporal, cingulate and parahippocampal gyri. Note that those regions are significantly different in cortical thickness between schizophrenic

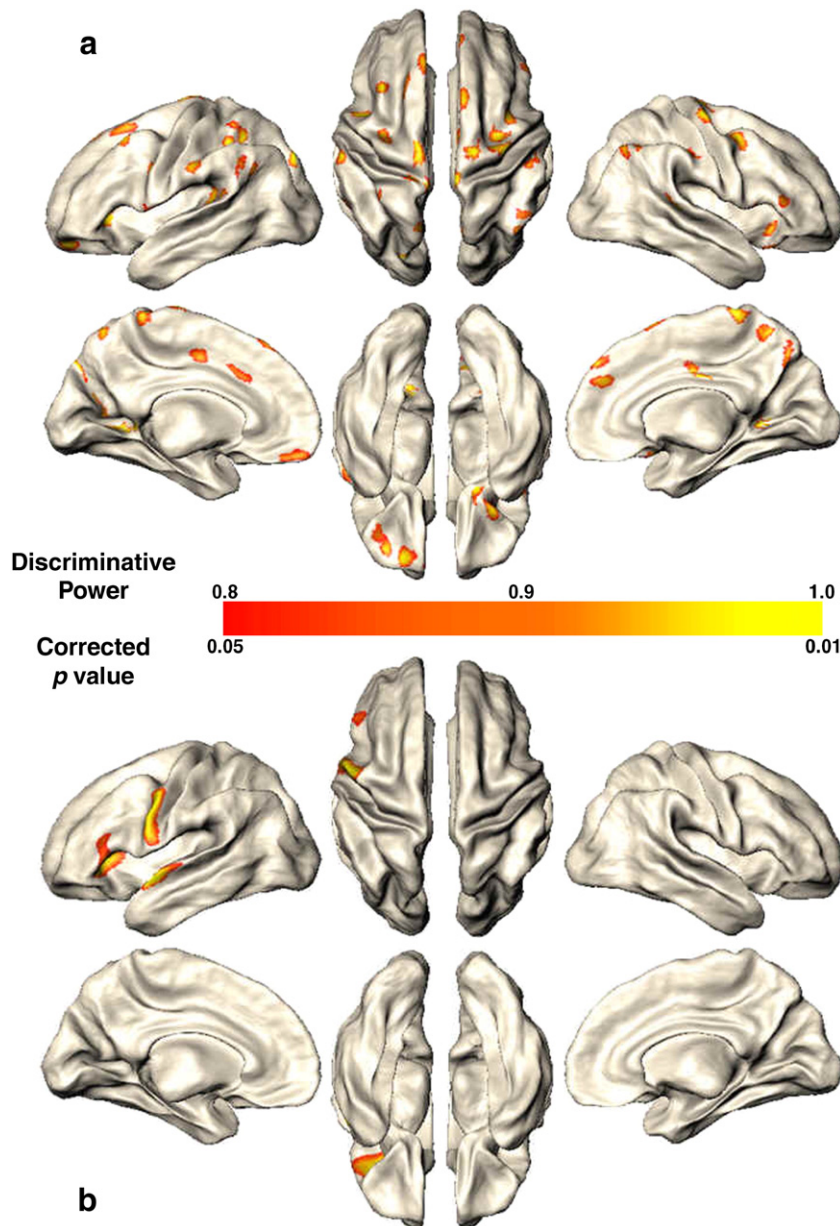


Fig. 3. The representative regions (a) and result of vertex-wise t -test (b) were displayed by color-coded mapping normalized discriminative power and corrected P value onto the mean cortical surface. Therefore, it can be said that there are significant differences of cortical thickness in these regions which contribute considerably to the classification.

patients and healthy controls. For the comparison with discriminative pattern, Fig. 3(b) shows the color-coded result of statistical t -test for the cortical thickness of patients and controls. There were statistically significant regions in the only left hemisphere, which included precentral, inferior frontal and superior temporal gyri. The spatial locations of significant regions were somewhat different with the discriminative patterns: the areas of the precentral and inferior frontal gyri were much more prominent and the significant location of superior temporal gyrus was a little more anterior in the t -map than the discriminative pattern. However, it is important to note that the voxel-by-voxel t -test shows the same anatomical labels as the discriminative pattern. This issue will be addressed in detail at the following section.

Discussion and conclusion

Principal component of cortical thickness as a feature for pattern classification

The original data are represented by fewer variables with minimal mean square error as a result of PCA, which reduces the dimensionality of the dataset. However, one of the limitations is that PCA only defines a single projection of the data. For more complex data like in this study, different clusters may require different projection directions. In addition, even if uncorrelated, the principal components might be highly statistically dependent. The other limitation of PCA is that the original data should have a linear or near-linear structure, to ensure the singularity of the data matrix. If

Table 3
Anatomical label and Talairach coordinate of representative region were obtained by *Talairach Daemon Client* software

Lobes	<i>x</i>	<i>y</i>	<i>z</i>	Location
Frontal	-18	52	-15	L superior frontal gyrus
	8	52	22	R superior frontal gyrus
	-30	47	-11	L middle frontal gyrus
	37	5	46	R middle frontal gyrus
	-31	31	0	L inferior frontal gyrus
	26	28	-13	R inferior frontal gyrus
	-7	50	-15	L medial frontal gyrus
	9	44	24	R medial frontal gyrus
	-33	-6	48	L precentral gyrus
Temporal	30	-13	62	R precentral gyrus
	-41	-32	14	L superior temporal gyrus
	60	-31	8	R superior temporal gyrus
Parietal	-56	-19	32	L postcentral gyrus
	8	-40	67	R postcentral gyrus
	-21	-79	39	L precuneus
	9	-64	45	R precuneus
Limbic	-11	-7	40	L cingulate gyrus
	8	-4	36	R cingulate gyrus
	-21	-61	12	L posterior cingulate gyrus
	-23	-50	2	L parahippocampal gyrus
	28	-50	4	R parahippocampal gyrus

It was defined from whole discriminative patterns according to value passing a threshold and the size of the cluster.

the data have a nonlinear structure as this study, the linear PCA may not be adequate in exploring the data. Although this simple realization has prompted the development of nonlinear alternatives to PCA, the purpose of this study is the identification of representative regions contributing to the classification through a discriminative pattern which is calculated by eigenvectors depending on selected principal components. Therefore, nonlinear PCA cannot be useful in this study. If the normative principal components are derived from the control group data matrix alone, it should be the best way to reduce dimensionality for classification. However, the results of such an analysis might be unstable because the number of control subjects (52) is small relative to the number of variables (4000–15,000 vertices). We applied PCA to the whole dataset in a single step, rather than constructing a new basis for each LOOCV trial based on individual training sets. Although it could be a basis for overtraining of classifier and a limitation of this study, Shen et al. (2004) mentioned that it was a simpler approach to minimize representation errors, which used principal components as features in the classification. A future direction of study will be involved in finding normative principal components of cortical thickness with larger samples and nonlinear techniques.

It is common to assess the performance of the rule for a selected subset by its LOOCV error, regardless of how the performance of the rule is assessed during the feature selection process. But, if it is calculated within the feature selection process, there is a selection bias in it when it is used as an estimate of the prediction error (Ambroise and McLachlan, 2002). External cross-validation should be undertaken subsequent to the feature selection process to correct for this selection bias. Alternatively, the bootstrap could be used so that we simulated using 50 random-selected validation set, which separated from the training set for the generalization. The results showed that using too few features was not able to separate the classes well while using excessively many features induced unnecessary noise (refer to Fig. 2(b)).

Feature selection is a machine learning method that reduces the number of statistical features in high-dimensional classification problems by finding subsets of features that are most relevant for discrimination. The method dramatically improves the generalization ability of a classifier in high-dimensional low sample size problems (Weston et al., 2000; Yushkevich et al., 2003). A number of methods use statistical classification to gain insight into the shape differences of biological objects in distinct classes of subjects. In particular, several classification studies have been conducted for elucidation of hippocampal shape abnormalities in schizophrenia (Csernansky et al., 1998; Golland et al., 2005; Shen et al., 2003, 2004). The hippocampus is a critical structure of the human limbic system involved in learning and memory processing. Although cortical thickness in schizophrenic patients has been examined using postmortem and *in vivo* imaging data, there has been no pattern classification study using features extracted from the entire cortical surface model. To our knowledge, this is the first study of classification of schizophrenic patients and healthy controls based on principal components of cortical thickness. In addition, we achieved the best mean accuracy of more than 90% for the classification using 50 random-selected validation set. It would be competitive with the best result in the classification studies for hippocampal shape (Csernansky et al., 1998; Golland et al., 2005; Shen et al., 2003, 2004). However, it should be borne in mind that the result was acceptable because this study aimed to examine a discriminative pattern, rather a classification itself.

This study showed that reordering of principal components associated with statistical tests was much more effective features of classification. In addition to showing that a high level of classification accuracy is possible, the results with reordering might be also instructive in determining how many dimensions are required to support perfect classification and how many seem excessive. Even if only cortical thickness were used for classification in this study, the other geometrical information from cortical surface model should be extracted and tested for classification. Mean curvature, Gaussian curvature, or average convexity, for example, will be particularly informative for ascertaining the major cortical folds, and these parameters may further improve the classification accuracy.

Cortical thickness abnormalities in schizophrenia

Most studies suggest that the pathophysiology of schizophrenia affects several spatially diffused cortical regions such as the lateral and medial temporal cortices, which might manifest as abnormalities in cortical thickness (Lawrie and Abukmeil, 1998; Shenton et al., 2001; Wright et al., 1999a,b, 2000). Although less frequently consistent, GM deficits of schizophrenia in other regions, including prefrontal and orbitofrontal regions of frontal cortices, and parietal cortices, are also observed (Shenton et al., 2001).

Our findings in discriminative patterns showed that there were significant differences of cortical thickness in the bilateral superior temporal and parahippocampal gyri. In particular, GM volume reduction in the parahippocampal gyrus involving both the medial temporal lobe and the superior temporal gyrus has a large percentage of replication in VBM and volumetric approaches (Shenton et al., 2001).

The representative regions in the frontal lobe were the bilateral superior, middle, inferior and medial frontal gyri and precentral gyrus. MRI findings of frontal lobe volume deficits in schizophrenia have been relatively inconsistent, with several studies

reporting reduced volume in schizophrenia and others reporting no significant differences (Goldstein et al., 1999; McCarley et al., 1999; Shenton et al., 2001). Although a few studies have evaluated subregions within the frontal lobe, Buchanan et al. (1998) reported volume reductions in prefrontal WM as well as right and left inferior gyri volume reduction in schizophrenia. Cortical volume reduction of precentral gyrus has been reported, which partially results in impaired motor functioning as well as abnormal physiological and primitive reflexes in schizophrenia (Ismail et al., 1998).

The regions in parietal lobe having significant discriminative power were the postcentral gyrus and the precuneus bilaterally. Although the parietal lobe is an important brain region for language processing, spatial working memory and attention, which are disturbed in schizophrenia, it has received very little attention (Shenton et al., 2001). Recently, several VBM studies have reported GM reduction of the left postcentral gyrus, but increase of the precuneus associated with better immediate verbal memory in patients (Antonova et al., 2005; Job et al., 2002).

Finally, other important regions worthy of investigation were the bilateral cingulate gyri. It is a pivotal component of brain networks directing affective and cognitive functions, and abnormalities of that region may be involved in the pathophysiology of schizophrenia. Although morphological studies of cingulate gyri have been few and inconsistent, Kopelman et al. (2005) reported that the schizophrenic patient group had a significantly larger left anterior cingulate gyrus volume relative to comparison subjects.

As described above, most representative regions having significantly discriminative power in this study were generally consistent with previous image-based or surface-based analysis, but quite focal. Although there were several reasons, PCA for dimension reduction should affect primarily those results. Generally, PCA extracts patterns explaining the greatest variance in the data. It cannot be guaranteed whether those patterns in the original space are continuous. In addition, the clusters of our result seem associated with network implicated in schizophrenia such as the frontotemporal network for processing auditory information and language (Gaser et al., 2004). The proposed method in this study permitted more precise description of the differences between groups so that it would be robust basis for diagnostic application.

Methodological considerations

Since most image processing algorithms, notably segmentation, are sensitive to noise, improvement of the signal to noise ratio would improve our algorithm. And, anisotropic diffusion (AD) filter, which was used for noise reduction in this study, was designed to operate on images consisting of homogeneous regions with slowly varying intensities, separated by well-defined boundaries. Since the bias field changes the contrast across anatomical edges throughout the image, preprocessing with the AD filter tends to suppress the edges more in regions of low bias than in regions of high bias (low bias is when bias is small and creates a large gray level distortion, high bias causes little or no gray level distortion). Guillemaud and Brady (1997) reported that the estimated bias field is essentially the same with or without prefiltering to remove noise. However, the final segmentation of the data is sensitive to noise, which involves determining the maximum likelihood tissue type for each voxel. Although it is beyond the scope of our study, the effect of the sequential order of bias field correction and AD filtering on the cortical thickness

would be necessary to investigate in the future. On the other hand, even if a wider space between the banks of a sulcus could lead to different behavior of AD filtering, it might be out of the question because the main point of this study was the cortical thickness between WM and GM surfaces and AD filtering would make a same effect on those surfaces. In addition, our previous volumetric study showed there was no significant effect of AD filtering on tissue classification (Yoon et al., 2003).

Generally, there might be several limitations associated with measuring cortical thickness from imaging data. For example, if the image contrast is reduced because of partial volume effects as intrinsic artifacts of MRI, GM boundaries might be irregular, leading to spatial inaccuracies in thickness value. However, cortical thickness measured at the nodes of a polygonal mesh rather than a voxel grid should have the advantage of providing a direct and accurate quantitative index of cortical morphology if the faithful reconstruction of surface models that represent the true cortical surface in both geometry and topology is guaranteed. Previously, we performed a quantitative cross-validation of the most prominent cortical surface reconstruction tools, including CLASP used in this study (Lee et al., 2006). Although there might be different viewpoints for validation, CLASP showed the best geometric and topological accuracy and mesh characteristics. As described above, the geometric accuracy of surface extraction is essential for the accurate measurement of morphological variables such as cortical thickness and cortical surface area. Therefore, it is reliable to use the cortical thickness values obtained by the methods used in this study as a feature for pattern classification.

It is notable that the univariate (voxel-by-voxel) statistical analysis, which is a hypothesis-driven approach, shows the same anatomical labels as the discriminative pattern generated from PCA and linear classifier, which is a purely data-driven method. The proposed method is able to identify unexpected sources of difference without any prior knowledge of data so that it can be seen as a hypothesis generating method. However, the representative regions of discriminative pattern and statistical *t*-test were somewhat different with each other (refer to Fig. 3). It might come from inevitable defects of feature selection for classification. For instance, the results of PCA might be somewhat unstable due to the restricted number of subject relative to the variables such as the number of vertex. And, if there is imperfection in the rearrangement of feature for classification, a certain principal component which included specific regions might be excluded from a classification. Both of them need to be tested in further studies.

Although PCA was applied to reduce dimensionality and make classification feasible in this study, its utility depends on whether meaningful biological interpretation can be described as different components, as PCA is an uninformed multivariate procedure. Wright et al. (1999b) conducted an exploratory analysis to identify supra-regional brain systems and to investigate whether abnormal brain architecture in schizophrenia manifests within one or more of these systems. Although they meaningfully interpreted components of regional brain measures extracted by PCA in terms of normal brain architecture and function, there was a critical limitation on automated cerebral parcellation due to the inter-individual variability in the anatomical boundaries of Brodmann regions. In a future study, we will attempt to overcome this kind of defect using a reconstruction of cortical surface model that is accurate in geometry and topology, and investigate the clinical meaning of principal components from cortical thickness measurements.

Although Davatzikos et al. (2005) were able to classify new individuals as schizophrenic or healthy with the reasonable accuracy, larger samples were necessary to replicate the results and establish regional group differences. Rather than classification itself, this study concerned to identify representative regions contributing to the classification through a discriminative pattern. Even if the discriminative pattern is less conservative than statistical *t*-test, this approach including the classifier interpretation in terms of the original features should be an important component of the data analysis in many applications where the statistical learning techniques are used to discover and study structural differences in the data. In addition, our classification results suggest that complex spatial patterns of morphometric characteristics can be used as an aid in the diagnosis of schizophrenia. The leave-one-out procedure is well established as a means for obtaining relatively unbiased estimates of the expected accuracy of the resulting classifier function, but for a small number of examples, the variance of this estimator might lead to a substantial difference between the estimated and the actual measurements. As the number of available examples increases, a more accurate estimation of the expected accuracy will be allowed. We also examined more effective feature selection schemes for improved classification accuracy. Although the training algorithm will produce a classifier that can be analyzed for discriminative pattern, the shape differences found in the training set are useful for understanding the phenomenon only if they accurately reflect the morphological differences in the entire population (Golland et al., 2005). As discriminative patterns derived from the classifier were consistent with a previous morphological analysis of schizophrenic patients, as described above, it can be said that the potential benefits of such diagnostic tools are improved by these findings.

Acknowledgment

This work was supported by the research fund of Hanyang University (HY-2004-N).

References

- Ambrose, C., McLachlan, G.J., 2002. Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc. Natl. Acad. Sci. U. S. A.* 99 (10), 6562–6566.
- Antonova, E., Kumari, V., Morris, R., Halari, R., Anilkumar, A., Mehrotra, R., Sharma, T., 2005. The relationship of structural alterations to cognitive deficits in schizophrenia: a voxel-based morphometry study. *Biol. Psychiatry* 58 (6), 457–467.
- APA, 1987. *Diagnostic and Statistical Manual of Mental Disorders*, 3rd ed., revised. American Psychiatric Press Inc., Washington, DC.
- Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. *NeuroImage* 11 (6 Pt. 1), 805–821.
- Buchanan, R.W., Vladar, K., Barta, P.E., Pearlson, G.D., 1998. Structural evaluation of the prefrontal cortex in schizophrenia. *Am. J. Psychiatry* 155 (8), 1049–1055.
- Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. *Data Mining Knowl. Discov.* 2 (2), 121–167.
- Burns, J., Job, D., Bastin, M.E., Whalley, H., Macgillivray, T., Johnstone, E.C., Lawrie, S.M., 2003. Structural disconnectivity in schizophrenia: a diffusion tensor magnetic resonance imaging study. *Br. J. Psychiatry* 182, 439–443.
- Cheverud, J.M., 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution* 36, 499–516.
- Chung, M.K., Worsley, K.J., Paus, T., Robbins, S., Evans, A.C., Taylor, J., Giedd, J.N., Rapoport, J.L., 2002. *Tensor-based Surface Morphometry*. University of Wisconsin, Madison.
- Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* 18 (2), 192–205.
- Cox, R.W., Hyde, J.S., 1997. Software tools for analysis and visualization of fMRI data. *NMR Biomed.* 10 (4–5), 171–178.
- Csemansky, J.G., Joshi, S., Wang, L., Haller, J.W., Gado, M., Miller, J.P., Grenander, U., Miller, M.I., 1998. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* 95 (19), 11406–11411.
- Davatzikos, C., Shen, D., Gur, R.C., Wu, X., Liu, D., Fan, Y., Hughett, P., Turetsky, B.I., Gur, R.E., 2005. Whole-brain morphometric study of schizophrenia revealing a spatially complex set of focal abnormalities. *Arch. Gen. Psychiatry* 62, 1218–1227.
- Duda, R.O., Hart, P.E., Stork, D., 2000. *Pattern Classification*, 2nd ed. John Wiley and Sons, New York.
- Fan, R.E., Chen, P.H., Lin, C.J., 2005. Working set selection using the second order information for training SVM. *J. Mach. Learn. Res.* 6, 1889–1918.
- First, M.B., Spitzer, R.L., Williams, J.B., Gibbon, M., 1997. *User's Guide for the Structured Clinical Interview for DSM-IV AXIS I Disorders (SCID-I): Clinician Version*. American Psychiatric Publishing, Inc., New York.
- Gaser, C., Nenadic, I., Volz, H.P., Buchel, C., Sauer, H., 2004. Neuroanatomy of “hearing voices”: a frontotemporal brain structural abnormality associated with auditory hallucinations in schizophrenia. *Cereb. Cortex* 14 (1), 91–96.
- Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15 (4), 870–878.
- Goldstein, J.M., Goodman, J.M., Seidman, L.J., Kennedy, D.N., Makris, N., Lee, H., Tourville, J., Caviness Jr., V.S., Faraone, S.V., Tsuang, M.T., 1999. Cortical abnormalities in schizophrenia identified by structural magnetic resonance imaging. *Arch. Gen. Psychiatry* 56 (6), 537–547.
- Golland, P., Grimson, W.E., Shenton, M.E., Kikinis, R., 2005. Detection and analysis of statistical differences in anatomical shape. *Med. Image Anal.* 9 (1), 69–86.
- Guillemaud, R., Brady, M., 1997. Estimating the bias field of MR images. *IEEE Trans. Med. Imag.* 16 (3), 238–251.
- Hughes, G.F., 1968. On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inf. Theory* 14, 55–63.
- Im, K., Lee, J.M., Lee, J., Shin, Y.W., Kim, I.Y., Kwon, J.S., Kim, S.I., 2006. Gender difference analysis of cortical thickness in healthy young adults with surface-based methods. *NeuroImage* 31 (1), 31–38.
- Ismail, B., Cantor-Graae, E., McNeil, T.F., 1998. Neurological abnormalities in schizophrenic patients and their siblings. *Am. J. Psychiatry* 155 (1), 84–89.
- Job, D.E., Whalley, H.C., McConnell, S., Glabus, M., Johnstone, E.C., Lawrie, S.M., 2002. Structural gray matter differences between first-episode schizophrenics and normal controls using voxel-based morphometry. *NeuroImage* 17 (2), 880–889.
- Kabani, N., Le Goualher, G., MacDonald, D., Evans, A.C., 2001. Measurement of cortical thickness using an automated 3-D algorithm: a validation study. *NeuroImage* 13 (2), 375–380.
- Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13 (2), 261–276.
- Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C., 2005. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *NeuroImage* 27 (1), 210–221.
- Kopelman, A., Andreasen, N.C., Nopoulos, P., 2005. Morphology of the anterior cingulate gyrus in patients with schizophrenia: relationship to typical neuroleptic exposure. *Am. J. Psychiatry* 162 (10), 1872–1878.
- Kuperberg, G.R., Broome, M.R., McGuire, P.K., David, A.S., Eddy, M.,

- Ozawa, F., Goff, D., West, W.C., Williams, S.C., van der Kouwe, A.J., 2003. Regionally localized thinning of the cerebral cortex in schizophrenia. *Arch. Gen. Psychiatry* 60 (9), 878–888.
- Lawrie, S.M., Abukmeil, S.S., 1998. Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies. *Br. J. Psychiatry* 172, 110–120.
- Lee, J.K., Lee, J.M., Kim, J.S., Kim, I.Y., Evans, A.C., Kim, S.I., 2006. A novel quantitative cross-validation of different cortical surface reconstruction algorithms using MRI phantom. *NeuroImage* 31 (2), 572–584.
- Lerch, J.P., Evans, A.C., 2005. Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage* 24 (1), 163–173.
- Lerch, J.P., Pruessner, J.C., Zijdenbos, A., Hampel, H., Teipel, S.J., Evans, A.C., 2005. Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy. *Cereb. Cortex* 15 (7), 995–1001.
- Lerch, J.P., Worsley, K., Shaw, W.P., Greenstein, D.K., Lenroot, R.K., Giedd, J., Evans, A.C., 2006. Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. *NeuroImage* 31 (3), 993–1003.
- MacDonald, D., Kabani, N., Avis, D., Evans, A.C., 2000. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. *NeuroImage* 12 (3), 340–356.
- McCarley, R.W., Wible, C.G., Frumin, M., Hirayasu, Y., Levitt, J.J., Fischer, I.A., Shenton, M.E., 1999. MRI anatomy of schizophrenia. *Biol. Psychiatry* 45 (9), 1099–1119.
- Narr, K.L., Bilder, R.M., Toga, A.W., Woods, R.P., Rex, D.E., Szeszko, P.R., Robinson, D., Sevy, S., Gunduz-Bruce, H., Wang, Y.P., 2005a. Mapping cortical thickness and gray matter concentration in first-episode schizophrenia. *Cereb. Cortex* 15 (6), 708–719.
- Narr, K.L., Toga, A.W., Szeszko, P., Thompson, P.M., Woods, R.P., Robinson, D., Sevy, S., Wang, Y., Schrock, K., Bilder, R.M., 2005b. Cortical thinning in cingulate and occipital cortices in first episode schizophrenia. *Biol. Psychiatry* 58 (1), 32–40.
- Olson, E.C., Miller, R.L., 1958. *Morphological Integration*. University of Chicago Press, Chicago, IL.
- Robbins, S., 2003. Anatomical standardization of the human brain in Euclidean 3-space and on the cortical 2-manifold. PhD thesis. School of Computer Science, McGill University, Canada.
- Robbins, S., Evans, A.C., Collins, D.L., Whitesides, S., 2004. Tuning and comparing spatial normalization methods. *Med. Image Anal.* 8 (3), 311–323.
- Selemon, L.D., 2004. Increased cortical neuronal density in schizophrenia. *Am. J. Psychiatry* 161 (9), 1564.
- Selemon, L.D., Rajkowska, G., Goldman-Rakic, P.S., 1995. Abnormally high neuronal density in the schizophrenic cortex. A morphometric analysis of prefrontal area 9 and occipital area 17. *Arch. Gen. Psychiatry* 52 (10), 805–818 (discussion 819–820).
- Shen, L., Ford, J., Makedon, F., Saykin, A., 2003. Hippocampal shape analysis: surface-based representation and classification. *Proc.-SPIE* 253–264.
- Shen, L., Ford, J., Makedon, F., Saykin, A., 2004. A surface-based approach for classification of 3D neuroanatomic structures. *Intell. Data Anal., Int. J.* 8 (6), 519–542.
- Shenton, M.E., Dickey, C.C., Frumin, M., McCarley, R.W., 2001. A review of MRI findings in schizophrenia. *Schizophr. Res.* 49 (1–2), 1–52.
- Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imag.* 17 (1), 87–97.
- Vapnik, V., 1998. *Statistical Learning Theory*. John Wiley and Sons, New York, NY.
- Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., Vapnik, V., 2000. Feature selection for SVMs. *Adv. Neural Inf. Process. Syst.* 13, 668–674.
- White, T., Andreassen, N.C., Nopoulos, P., Magnotta, V., 2003. Gyrfication abnormalities in childhood- and adolescent-onset schizophrenia. *Biol. Psychiatry* 54 (4), 418–426.
- Wiegand, L.C., Warfield, S.K., Levitt, J.J., Hirayasu, Y., Salisbury, D.F., Heckers, S., Dickey, C.C., Kikinis, R., Jolesz, F.A., McCarley, R.W., 2004. Prefrontal cortical thickness in first-episode psychosis: a magnetic resonance imaging study. *Biol. Psychiatry* 55 (2), 131–140.
- Wright, I.C., Ellison, Z.R., Sharma, T., Friston, K.J., Murray, R.M., McGuire, P.K., 1999a. Mapping of grey matter changes in schizophrenia. *Schizophr. Res.* 35 (1), 1–14.
- Wright, I.C., Sharma, T., Ellison, Z.R., McGuire, P.K., Friston, K.J., Brammer, M.J., Murray, R.M., Bullmore, E.T., 1999b. Supra-regional brain systems and the neuropathology of schizophrenia. *Cereb. Cortex* 9 (4), 366–378.
- Wright, I.C., Rabe-Hesketh, S., Mellers, J., Bullmore, E.T., 2000. Testing for laterality differences in regional brain volumes. *Arch. Gen. Psychiatry* 57 (5), 511–512.
- Yoon, U., Lee, J.M., Kim, J.J., Lee, S.M., Kim, I.Y., Kwon, J.S., Kim, S.I., 2003. Modified magnetic resonance image based parcellation method for cerebral cortex using successive fuzzy clustering and boundary detection. *Ann. Biomed. Eng.* 31 (4), 441–447.
- Yoon, U., Lee, J.M., Koo, B.B., Shin, Y.W., Lee, K.J., Kim, I.Y., Kwon, J.S., Kim, S.I., 2005. Quantitative analysis of group-specific brain tissue probability map for schizophrenic patients. *NeuroImage* 26 (2), 502–512.
- Yushkevich, P., Joshi, S., Pizer, S.M., Csernansky, J.G., Wang, L.E., 2003. Feature selection for shape-based classification of biological objects. *Lect. Notes Comput. Sci.* 2732, 114–125.
- Zijdenbos, A.P., Forghani, R., Evans, A.C., 2002. Automatic “pipeline” analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Trans. Med. Imag.* 21 (10), 1280–1291.