Estimating Cluster Overlap on Manifolds and its Application to Neuropsychiatric Disorders

Peng Wang [1], Christian Kohler [2], Ragini Verma [1]
[1] Department of Radiology
University of Pennsylvania
Philadelphia, PA 19104
{wpeng@ieee.org, ragini.verma@uphs.upenn.edu }

[2] Department of Psychiatry
University of Pennsylvania
Philadelphia, PA 19104
kohler@bbl.med.upenn.edu

Abstract

Although it is usually assumed in many pattern recognition problems that different patterns are distinguishable, some patterns may have inseparable overlap. For example, some facial expressions involve subtle muscle movements, and are difficult to separate from other expressions or neutral faces. In this paper, we consider such overlapped patterns as "clusters", and present a novel method to quantify cluster overlap based on the Bayes error estimation on manifolds. Our method first applies a manifold learning method, ISOMAP, to discover the intrinsic structure of data, and then measures the overlap of different clusters using the k-NN Bayes error estimation on the learned manifolds. Due to the ISOMAP's capability of preserving geodesic distances and k-NN's localized estimation, the method can provide an accurate measure of the overlap between clusters, as demonstrated by our simulation experiments. The method is further applied for an analysis of a specific type of facial expression impairment in schizophrenia, i.e.,"flat effect", which refers to a severe reduction in emotional expressiveness. In this study, we capture facial expressions of individuals, and quantify their expression flatness by estimating overlap between different facial expressions. The experimental results show that the patient group has much larger facial expression overlap than the control group, and demonstrate that the flat affect is an important symptom in diagnosing schizophrenia patients.

1. Introduction

In pattern classification problems, it is usually assumed that different patterns can be well separated. However, real data always has overlap, which will cause classification errors [3]. Such overlap could be generated from noisy measurements, or from continuous events. For example, some facial expressions involve subtle and continuous muscle movements, and are difficult to be separated from each other or neutral faces. Another example is the face pose estimation, where the face pose change is also a continuous process. There are only approximate, not absolute, class boundaries to separate faces into frontal, left and right views. Therefore, we consider such overlapped patterns as "clusters". Due to the uncertainty of patterns, it is difficult to quantitatively evaluate the cluster overlap, especially when the data has high dimensionality. In this paper, we present a novel method to estimate the overlap between different clusters based on Bayes error estimation on manifolds.

Given a dataset whose elements are generated from overlapped clusters, we first apply a manifold learning method, ISOMAP [13], to discover the intrinsic structure of the data. The advantages of ISOMAP manifold learning are that it can reduce the data dimensionality, and recover the intrinsic structure of the data. Then, based on the learned manifold, a k-NN based method of estimating Bayes error is applied to quantify the cluster overlap. The benefit of estimating cluster overlap on manifolds is that ISOMAP can reduce data dimensionality and preserve neighborhood information as well, thus facilitating the distance measurement during k-NN based Bayes error estimation.

This method is first demonstrated by experiments using simulated data, and then is successfully applied for the analysis of a specific type of facial expression impairment in schizophrenia, "flat effect", which refers to a severe reduction in emotional expressiveness. In this application, facial features are first extracted through video tracking. For each individual, the extracted features are embedded into an ISOMAP manifold. The overlap between different expressions are measured on the manifold, to describe the flatness of facial expressions for the individual. By comparing the expression overlap between the groups of schizophrenia patients and healthy controls, our method demonstrates that the patient group has much larger overlap between ex-

pressions, thus quantifying the presence of facial expression flatness in schizophrenia and thereby aiding in diagnosis, prognosis or studying treatment effects.

The rest of the paper is organized as follows. In Section 2, the related work is reviewed. The method of measuring cluster overlap using manifold is introduced in Section 3. Section 4 introduces its application to the analysis of facial expression impairments in a neuropsychiatric disorder. The paper concludes in Section 5.

2. Related Work

This section reviews some related work, including commonly used linear and nonlinear component analysis, and the cluster overlap estimation. The linear component analysis methods, including the classic Principal Component Analysis (PCA) and Fisher Linear Discriminant Analysis (LDA), have been widely applied to solve many computer vision and pattern recognition problems. A well-known example is face recognition [8, 17]. However, both PCA and LDA have the assumption that each class follows Gaussian distribution. Nonparametric discriminant analysis (NDA) [4] can avoid such limitations, by computing scatter matrices in neighborhoods instead of globally. The linear component analysis using the matrix representation has been generalized to the tensor space that can provide multi-linear analysis [14]. However, such linear methods will fail for the cases where the data has a non-linear structure. The efforts to generalize the linear component analysis can be categorized into two types: kernel-based methods, and the nonlinear manifold learning algorithms. The kernel based methods, such as Kernel PCA (KPCA) [1] and Kernel LDA (KLDA) [11], implicitly map the data to a higher dimensional space through kernel functions [7]. Nonlinear manifold learning methods, such as ISOMAP [13], and LLE [10], aim at recovering the nonlinear structures of data by using their neighboring relationship. Such manifold methods are superior to the linear methods, in that they can handle problems with complicated nonlinear distributions.

In pattern classification, the features obtained from the component analysis are then used to design classifiers for distinguishing different patterns. The classification errors due to the noise and the underlying class overlap are usually measured empirically, such as through cross-validation [3]. However, such empirical validation largely depends on the feature selection and classifier design, and cannot reveal the underlying overlap between classes. Fukunaga et al. use Parzen windows and k-NN to estimate the Bayes error rates [5]. It has been shown that the error rate estimated based on the k-NN rule has upper bound of twice the true Bayes error [3]. The accuracy of k-NN estimation depends on the selection of the window function and its parameters. And furthermore, when the data has high dimensionality, the k-NN based method needs more computation and more data samples to locate nearest neighbors in the high dimensional space. Our method combines the merits of manifold learning and k-NN Bayes error estimation to measure the cluster overlap. The ISOMAP manifold learning method is used to reduce data dimension, remove noise, and preserve the geodesic distances for accurately searching nearest neighbors. Based on learned manifolds, the k-NN based method can then provide a reasonably accurate estimation of the underlying cluster overlap.

3. Manifold Based Analysis of Cluster Overlap

Given a set of data generated from m different patterns, which is denoted as $C_i, i = 1, ..., m$, the data can be assumed to lie in a manifold that characterizes different patterns. Due to noise and the possible underlying overlap, the data from different patterns also has overlap on a manifold, thus forming "clusters". In this section, we first briefly review the ISOMAP learning method that will be used in our method, and then present the method to estimate cluster overlap on the manifold.

3.1. ISOMAP Manifold Learning

In this method, we apply the ISOMAP manifold learning algorithm because it can preserve the geodesic distances of the original data. Such property allows us to measure the true "geodesic" distances on the learned ISOMAP manifold that has lower dimensionality. In ISOMAP, a shortest-path graph is first constructed based on local neighbors, and then the global nonlinear structure is recovered by applying multi-dimensional scaling (MDS) on the shortest distances. The ISOMAP learning method is summarized in Table 1.

3.2. A Probabilistic k-NN Model of Estimating Overlap

We present a probabilistic method to estimate Bayes error between different clusters as the measurement of their overlap, using the learned manifold. The probabilistic method is based on the k-NN estimation of Bayes error [4, 3]. The Bayes error is the minimal error rate that a probabilistic classifier can achieve based on the Bayes rule, while the k-NN estimation has an upper bound that is twice of the true Bayes error [3]. The k-NN method is also known to be sensitive to the selection of parameters of window functions, especially for high dimensional data. Our method is different from the k-NN estimation presented in [4], in that our method defines the nearest neighbors as obtained from the manifold learning, directly estimates the posterior probability of clusters, and automatically selects the parameter of the window function.

To estimate the Bayes error in a manifold, we first define the posterior probability of a cluster given a data point x, as

Table 1. ISOMAP algorithm

1. Input: the data set $\Omega = \{x_i, i = 1, ..., n\}$
2. Compute the pair-wise distance $d(x_i, x_j)$ between data points x_i and x_j, where d is the distance metric, usually L_2 norm.
3. Search the nearest neighbors *NN*(x_i), for each data point x_i in the set Ω. The neighbors can be found using one of the following two methods:
   (a) k-nearest neighbors (k-NN): finding k closet neighbors in Ω.
   (b) ϵ-nearest neighbors (ϵ-NN): $NN(x_i) = \{x_j | d(x_i, x_j) < \epsilon, i \neq j, x_j \in \Omega\}$
4. Construct a graph $G = \{V, E\}$ by connecting each data point with its neighbors. V represents a set of nodes in the graph. Each node represents a data point. E is a set of links between nodes, with the weights set as the distances between nodes.
5. Compute the shortest path distances from the graph G, and obtain a shortest distance matrix D.
6. Apply the multi-dimensional scaling (MDS) to D, and embed the original data points to the lower dimensional ISOMAP subspace.

Eqn. (1):

$$P(C_i|x) = \int_{y \in C_i} p(y|x)dy \tag{1}$$

where C_i is the cluster label, and y is the data from cluster C_i.

We assume the $p(y|x)$ as a Gaussian distribution $G(y; x, \sigma)$ with the mean of x and standard deviation of σ, as shown in Eqn. (2):

$$p(y|x) = G(y; x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\{\frac{-d^2(x, y)}{2\sigma^2}\} \tag{2}$$

$p(y|x)$ defined in Eqn. (2) is actually a window function that applies a kernel to each data point. Please note that there is only one parameter σ in this function. The σ will be automatically estimated from nearest neighbors in our method, as explained later. The class probability $P(C_i|x)$ can then be directly estimated based on Eqns. (1) and (2). Since the $d(x, y)$ is large when y is far from data x , we only use the data in the neighborhood of x to estimate $P(C_i|x)$ for reducing computational complexity, as shown in Eqn. (3).

$$\begin{aligned} P(C_i|x) &\approx \int_{y \in C_i, y \in NN_x} p(y|x)dy \\ &\propto \sum_{y \in C_i, y \in NN_x} p(y|x) \end{aligned} \tag{3}$$

In Eqn (3), NN_x is the neighborhood of x, and y are the nearest neighbors of x from cluster C_i. In our method, the neighborhood is defined on the manifold subspace after ISOMAP embedding, instead of the original data space. Assuming that the data x after embedding is $F(x)$, the geodesic distance $d_g(x, y)$ and the neighborhood NN_x after the embedding are defined as:

$$\begin{aligned} d_g(x, y) &= ||F(x) - F(y)||_2 \\ NN_x &= \{y | d_g(x, y) < \epsilon_x\} \end{aligned} \tag{4}$$

and

$$p(y|x) = \frac{1}{\sqrt{2\pi}\sigma} exp\{\frac{-d_g^2(x, y)}{2\sigma^2}\} \tag{5}$$

In Eqn. (4), ϵ_x decides the size of neighborhood. If ϵ_x is a fixed value for all x, the neighborhood is the ϵ-NN. If ϵ_x is varied to guarantee k nearest neighbors for each data, NN_x equals to the k-NN.

According to the definition, the Bayes error rate is estimated at each data point x, as Eqn. (7):

$$P(Error|x) = 1 - \frac{\max_i \{\lambda_i P(C_i|x)\}}{\sum_j \lambda_j P(C_j|x)} \tag{6}$$

where λ_i is the risk for misclassifying cluster C_i. For simplicity, we usually assume the equal risk for each class, i.e., $\forall i, j, \lambda_i = \lambda_j$. Therefore, by the definition, the pairwise Bayes error $P_{(C_i, C_j)}(Error)$ between two clusters C_i and C_j, is estimated as:

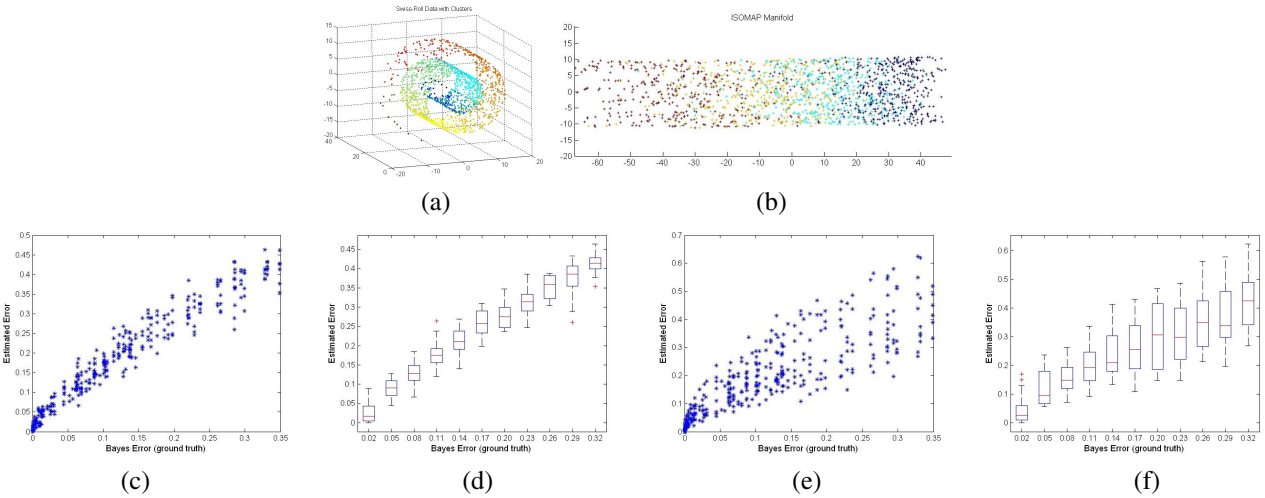
$$P_{(C_i, C_j)}(Error) = \sum_{x \in \{C_i, C_j\}} P(Error|x)P(x)dx \tag{7}$$

where $P(x)$ is usually assumed uniformly distributed in the data set. In what follows, we use the estimated Bayes error as the measurement of cluster overlap, defined as $Z^{i,j} = P_{(C_i, C_j)}(Error)$.

We go back to the problem of automatically determining the parameter σ in $p(y|x)$ from the manifold embedding. As the second order moment of a Gaussian distribution, the σ is estimated as Eqn. (8):

$$\sigma^2 = \underset{x, y \in NN_x}{E} \{d_g^2(x, y)\} \tag{8}$$

Again, the estimation is performed in the neighborhood obtained from ISOMAP manifolds. In summary, our estimation method fully utilizes the benefits of manifold learning algorithm, through reducing the data dimensionality and preserving geodesic distances in the neighborhood.



(a) (b) (c) (d) (e) (f)

Figure 1. Validation of overlap measurements using Swiss Roll data: (a) the Swiss Roll data with its parameter, i.e., the rolling angle, sampled at clusters; (b) ISOMAP embedding of the data at two dimensional space. Each color represents one cluster; (c) comparing the ground truth with the estimated Bayes error using the presented method; (d) a box and whisker plot of the results in (c); (e) comparing the ground truth with the estimated Bayes error using the k-NN based method [5] in the original data space; (f) a box and whisker plot of the results in (e)

3.3. Validation on Simulation Data

We simulate a set of data to validate the algorithm. Previous studies [13, 10] uniformly sample a Swiss Roll manifold for validation, while in this experiment, we generate a set of data on the Swiss Roll manifold by sampling its parameter space to form clusters with different levels of overlap. Assuming that there are n clusters in the data, the data of each cluster is simulated from a Gaussian distribution, $p_i(t) = G(t; \mu_i, \sigma_i), i = 1, ..., n$, where t is the parameter of the Swiss Roll manifold , and μ_i and σ_i are the mean and standard deviation of the i-th cluster. From the simulation processing, we know the ground truth, i.e., the Bayes error rate between two Gaussian distributions. For simplicity, all the clusters are assigned the same the priors, and the same σ_i. However, the center of each cluster is randomly varied, thus yielding different Bayes errors as the ground truth for comparison with the estimated overlap. The Swiss Roll data in the original data space is illustrated in Figure 1.(a). The ISOMAP embedding of one set of Swiss Roll data is shown in Figure 1.(b), where different colors represent different clusters. In the experiments, we use the same parameter as in [13], i.e., $\epsilon = 5$. It clearly shows that ISOMAP can recover the clusters on the original Swiss Roll data, and that there is an overlap between different clusters.

The simulated data after ISOMAP embedding is then used to estimate Bayes errors. The estimated cluster overlap is then compared with the ground truth, as shown in Figure 1.(c) and (d). It is found that the estimation provides an accurate estimation of the true Bayes error. The linear correlation coefficient between the estimation and the ground truth is around 0.97 for our results. Therefore it demonstrates that we can use this method to quantify the cluster overlap. For comparison, the original k-NN based method [5] is applied in the original data space. The estimation results are shown in Figure 1.(e) and (f), where we can see that the estimation results are less stable, especially for data sets that have overlap greater than 10%, demonstrating the benefits of measuring cluster overlap on the ISOMAP manifold.

4. Application: Measuring Flat Facial Expressions in Schizophrenia

In this section, we apply the presented method of measuring cluster overlap to the study of a population with the neuropsychiatric disorder of schizophrenia that manifests as flatness (lack or low intensity) of facial expressions. We use videos that contain different facial expressions of individuals for the analysis. By measuring the expression overlap, our method can quantify flat affect in schizophrenia. The background is introduced in Section 4.1. Section 4.2 presents the feature extraction in video for the analysis. The experimental results are summarized in Section 4.3.

4.1. Background

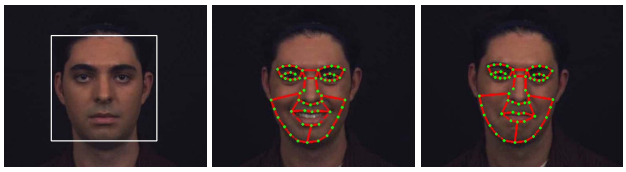
Facial expressions have been widely used in clinical research to study the affective and cognitive states, and psychopathology of an individual. Specifically, the expression analysis has played a major role in the study of schizophrenia which is a neuropsychiatric disorder characterized by deficits in emotional expressiveness [12]. One type of facial expression impairment that is often demonstrated by patients with schizophrenia is "flat affect", i.e., a severe reduction in emotional expressiveness. Examples are shown

in Figure.3, where a patient failed to express different emotions. In this study, the video data of 12 patients with schizophrenia and 12 healthy controls has been acquired under the supervision of psychiatrists. During the interviews with participants, they are individually guided through vignettes that are provided by the participants themselves and describe a situation in their life pertaining to each emotion. The vignettes are recounted back to the participants by the psychiatrist, to elicit evoked expressions. Videos captured during the interview include segments of different expressions and the neutral expression as well, and their time stamps of expressions are recorded for the later analysis.

Clinicians currently rely on manual and subjective methods of rating expressions, and the clinical research in schizophrenia has focused on the perception and recognition capabilities of the patients, and not so much on the way in which patients express emotions differently from healthy controls [12, 6]. Many automated facial expression analysis methods have been presented [9], with most of them focusing on facial expression recognition. None of them have been applied in clinical research to study the impaired facial expression in schizophrenia. In this paper, we apply our method of measuring the cluster overlap to quantify flat affect in schizophrenia patients using videos of facial expression.

4.2. Facial Feature Extraction in Video

Manually labeling fiducial points to analyze facial expressions is usually very time-consuming, and subjective to the person who labeled the face. Especially in our study, each video could contain more than 10,000 frames. Therefore it is a formidable task to manually mark all the fiducial points through the videos. To automate the process, we first detect the face and fiducial points in the first frame of the video, and then track the fiducial points at all the remaining frames. Our method applies an AdaBoost based algorithm [15] to detect frontal and near-frontal faces. Figure 2.(a) shows a face detection result at the first frame of a video. Inside the detected faces, some important fiducial points, as shown in 2.(b), are defined to characterize the facial expression changes. An active appearance model (AAM) [2] is used to automatically detect and track the fiducial points. Figure 2.(b) and (c) shows the detected and tracked fiducial landmarks in a video.



(a) (b) (c)

Figure 2. Fiducial landmarks detection and tracking. (a) a face detected at the first frame; (b)(c) the landmarks tracked at different frames

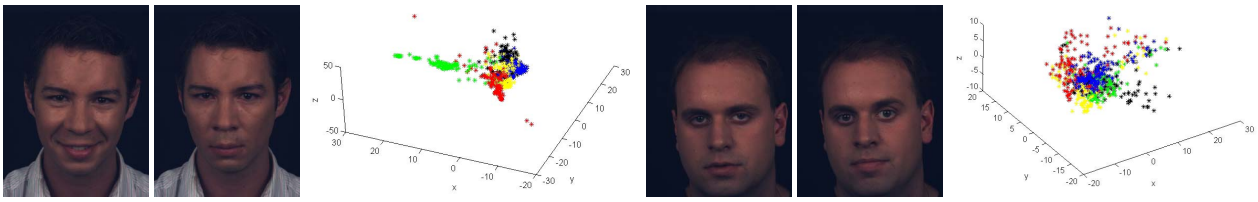
Based on the fiducial landmarks, geometric features are extracted for characterizing facial changes [16]. The geometric features describe the area changes of facial regions, and some facial actions that are known to be closely related with expressions, such as eye opening, mouth opening and closing. As the result, the extracted feature vector has 37 elements. Features of each participant are normalized by his/her features at neutral faces. Finally, each feature vector is normalized by subtracting its mean from it and dividing it by its standard deviation, for the following analysis on the manifold.

4.3. Measuring Flat Affect in Schizophrenia

For the analysis of the flatness of facial expressions in patients with schizophrenia, the features extracted from videos are embedded into ISOMAP manifolds. Our method builds an individual manifold for each participant, and the individual manifold will be used for subsequent overlap measurements. Since each video contains on average 10,000 frames, we downsample each video to around 2,000 data points, making it suitable for ISOMAP based learning. Our method uses k-NN during ISOMAP learning, where k is set as around 20. However, our experiments show that the results are robust to the parameter changes. Figure 3 show the example images and the ISOMAP manifolds of two participants, including one control and one patient. From this example, we clearly observe that there is more overlap in the schizophrenia patient than the healthy control.

After ISOMAP embedding, we measure the overlap among different expressions, including happiness, sadness, anger, fear, and neutral expressions. The pair-wise overlap $Z^{i,j} = P_{(C_i,C_j)}(Error)$ between the i-th expression and j-th expression are measured. We then define the minimal overlap between one expression to other expressions as a "flatness index" at all expressions, i.e. $Z^i = \underset{j \neq i}{min}\{Z^{i,j}\}$. A larger flatness index means that it is less possible to distinguish this expression from at least one of the other facial expressions. Although the length of each persons' expressions may vary, our methods uniformly normalize the weights of each data point, i.e., $P(x)$ in Eqn. (7), according to the cluster size; therefore the comparison of cluster overlap is not affected by cluster size.

The method is applied to each participant from both the healthy control group and the patient group. The results are summarized in Table 2. In this table, we show the mean flatness index for each group, at happy, sad, anger, fear and neutral expressions respectively. We also use the flatness index from all the participants of two groups for a t-test, to evaluate the significance level of the group difference. It is observed that the mean flatness index is much smaller in the healthy controls than in the patients, and the p-values from the t-test demonstrate that the group differences are



(a) (b) (c) (d) (e) (f)

Figure 3. Examples of ISOMAP manifold for two participants. (a)(b): happy and sad faces of a healthy control. (c) the ISOMAP manifold of the healthy control. Different colors represent data from different expressions; (d)(e): anger and fear faces of a patient; (f) the ISOMAP manifold of the patient. Visual comparison reveals that the patient expressions have a greater overlap than the controls, whose expression embedding is more clustered

significant, at all expressions. Among all the flatness measurements, we are especially interested in the flatness index of the neutral expression. Since the flatness index at neutral expression refers to the overlap between the emotional faces of an individual and his/her non-emotional faces, it directly measures the "flat affect". The flatness index at other expressions actually measures a mixture of flat affect and another type of impairment, "inappropriate affect", which refers to inappropriate expressions for intended emotions. Using the estimated overlap to measure inappropriate affect needs further research. In summary, by measuring the overlap among different expressions and non-emotional faces, our method quantitatively validates the clinical claim that the flat affect is an important symptom characterizing schizophrenia patients.

Table 2. Comparison of Expression Overlap between schizophrenia patients and controls using Flatness Index

Mean of flatness index	Happiness	Sadness	Anger	Fear	Neutral
Controls	0.057	0.091	0.051	0.030	0.093
Patients	0.263	0.340	0.310	0.239	0.241
P-Value of t-test	0.007	0.002	0.002	0.004	0.014

5. Conclusion

This paper presents a novel method to estimate the cluster overlap on manifolds. Our method first embeds the original data into ISOMAP manifolds to reduce the data dimensionality and recover its underlying structure. We then estimate the cluster overlap, by applying k-NN based estimation on the manifold. This method is applied to a clinical study to quantify a specific impairment, the "flat affect" of facial expressions, in schizophrenia patients. Our future work will correlate our method with clinical measures of flat affect, and will also apply this method to the feature extraction in general pattern recognition problems.

Acknowledgement

We would like to acknowledge support from NIH grants 1R01MH73174-01 (for method development) and R01-MH060722 (for data acquisition).

References

[1] Alexander Smola Bernhard Schlkopf and Klaus-Robert Mller, *Nonlinear component analysis as a kernel eigenvalue problem*, Neural Computation **10** (1998), 1299–1319.

[2] T.F. Cootes, G.J.Edwards, and C.J. Taylor, *Active appearance models*, IEEE. Trans. on Pattern Analysis and Machine Intelligence **23** (2001), no. 6, 681–685.

[3] Richard O. Duda, P.E.Hart, and David G. Stork, *Pattern classification*, second ed., John Willey Sons, 2000.

[4] K. Fukunaga, *Nonparametric discriminant analysis*, IEEE Transactions on Pattern Analysis and Machine Intelligence **5** (1984), no. 1, 671–678.

[5] K. Fukunaga and D. M. Hummels, *Bayes error estimation using parzen and k-nn procedures*, IEEE Trans. Pattern Analysis and Machine Intelligence **9** (1987), no. 5, 634–643.

[6] R. E. Gur, C. G. Kohler, J. D. Ragland, S. J. Siegel, K. Lesko, W. B. Bilker, and R. C. Gur, *Flat affect in schizophrenia: relation to edition processing and neurocoginitive measures*, Schizophrenia Bulletin **32** (2006), no. 2, 279–287.

[7] G. Rtsch K. Tsuda K.-R. Mller, S. Mika and B. Schlkopf, *An introduction to kernel-based learning algorithms*, IEEE Neural Networks **12** (2001), no. 2, 181–201.

[8] Baback Moghaddam and Alex Pentland, *Probabilistic visual learning for object representation*, IEEE Trans. on Pattern Analysis and Machine Intelligence **19** (1997), no. 7, 696–710.

[9] M. Pantic and L.J.M. Rothkrantz, *Automatic analysis of facial expressions: the state of the art*, IEEE. Trans. on Pattern Analysis and Machine Intelligence **22** (2000), no. 12, 1424–1445.

[10] S. Roweis and L. Saul, *Nonlinear dimensionality reduction by locally linear embedding*, Science **290** (2000), no. 5500, 2323–2326.

[11] J. Weston B. Schlkopf S. Mika, G. Rtsch and K.-R. Mller, *Fisher discriminant analysis with kernels*, Neural Networks for Signal Processing, 1999.

[12] J. E. Salem and A. M. Kring, *More evidence for generalized poor performance in facial emotion perception in schizophrenia*, Journal of Abnormal Psychology **105** (1996), no. 3, 480–483.

[13] J. B. Tenenbaum, V. de Silva, and J. C. Langford, *A global geometric framework for nonlinear dimensionality reduction*, Science **290** (2000), no. 5500, 2319–2323.

[14] M. A. O. Vasilescu and D. Terzopoulos, *Multilinear analysis of image ensembles: Tensorfaces*, ECCV, 2002.

[15] P. Viola and M. Jones, *Robust real-time object detection*, International Journal of Computer Vision **57** (2004), no. 2, 137–154.

[16] Peng Wang, Christian Kohler, Fred Barrett, Raquel Gur, Ruben Gur, and Ragini Verma, *Quantifying facial expression abnormality in schizophrenia by combining 2D and 3D features*, CVPR, 2007.

[17] Wenyi Zhao, *Discriminant component analysis for face recognition*, ICPR, vol. 2, 2000.