2010 Fourth International Conference on Network and System Security

Face Gender Recognition Based on 2D Principal Component Analysis and Support Vector Machine

Len Bui, Dat Tran, Xu Huang and Girija Chetty
University of Canberra, Faculty of Information Sciences and Engineering, Australia
{len.bui, dat.tran, xu.huang, girija.chetty}@canberra.edu.au

***Abstract*— This paper presents a novel method for solving face gender recognition problem. This method employs 2D Principal Component Analysis, one of the prominent methods for extracting feature vectors, and Support Vector Machine, the most powerful discriminative method for classification. Experiments for the proposed approach have been conducted on FERET data set and the results show that the proposed method could improve the classification rates.**

I. INTRODUCTION

Human faces contain a lot of important biometrics information. One of them is about sexual category. Therefore, advances in facial expression recognition would be meaningful to the development of biometrics authentication systems.

Since Matthew Turk and Alex Pentland [1] used Principal Component Analysis (PCA) to deal the face recognition problem, it has become the major mathematical tool to extract feature vectors. Nevertheless, PCA could not capture even the simplest invariance unless this information is explicitly provided in the training data. To deal with the problem, some researchers proposed other approaches. For example, Wiskott et al. [2] suggested a technique known as elastic bunch graph matching. Bartlett et al. [3] proposed using ICA for face representation and reported that its results were better than PCA. Ming-Hsuan Yang [4] suggested Kernel PCA for face feature extraction and recognition and described that his method outperformed the classical method. However, the performance costs of ICA and Kernel PCA are higher than PCA.

To solve these problems, Jian Yang [5] proposed a new method called 2D Principal Component Analysis (2DPCA). In conventional PCA, face images have been represented in vectors by some technique like concatenation. As opposed to PCA, 2DPCA represent face images by using matrices or 2D images instead of vectors (Fig. 1). Clearly, using 2D images directly is quite simple and local information of the original images can be preserved sufficiently, which may bring more important features for facial expression classification.

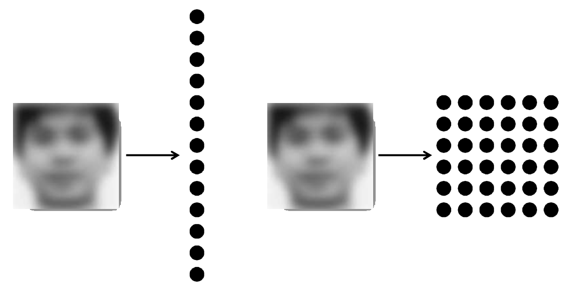

Figure 1. Image presentations in PCA and 2DPCA

In 1995, Vapnik and Cortes [6] presented the foundations for Support Vector Machine (SVM). Since then, it has become the prominent algorithm to solve problems in pattern classification and regression. The basic idea behind SVM is finding the optimal linear hyperplane such that the expected classification error for future test samples is minimized, i.e., good generalization performance. Obviously, the goal of all classifiers or regressors is not to get the lowest training error. For example, a k-NN classifier can achieve the accuracy rate 100% with k=1. However, in practice, it is the worst classifier because it has high structural risk.

They suggested the formula testing error = training error + risk of model (Fig. 2). To achieve the goal to get the lowest testing error, they proposed the structural risk minimization inductive principle. It means that a discriminative function that classifies the training data accurately and which belongs to a set of functions with the lowest VC dimension will generalize best regardless of the dimensionality of the input space. Based on this principle, an optimal linear discriminative function has been found. For linearly non-separable data, SVM map the input to a high dimensional feature space where a linear hyperplane can be found. Although there is no warranty that a

978-0-7695-4159-4/10 $26.00 © 2010 IEEE
DOI 10.1109/NSS.2010.19

IEEE computer society

linear solution will always exist in the high dimensional space, it is able to find effective solutions in practice. To deal with the face gender classification, many researchers [7-10] have applied SVM in their studies and stated that the experiment results are very positive. In our research, we have combined the powerful of each methods, 2DPCA and SVM, to solve the problem.

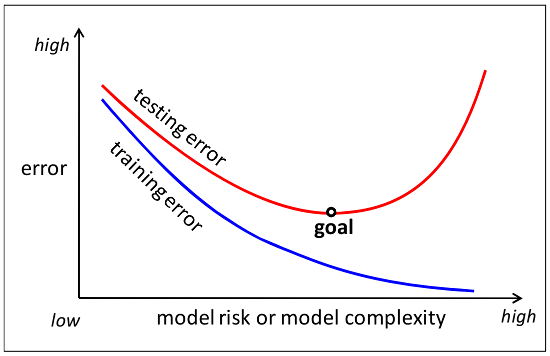

Figure 2. Curves of testing error and training error

The remaining sections of our paper will discuss the implementation of our face gender classification system, related theory, and experiments. Section 2 gives details of 2DPCA. Section 3 discusses how to use SVM for face gender classification. In Section 4, we will describe the implementation and experiments. Finally, we will conclude in Section 5.

II. Two dimension Principal Component Analysis

A. *Face model construction*

Our face model is built using a mathematical method called 2DPCA as mentioned above. As opposed to the conventional 2DPCA, we propose a novel 2DPCA with the weights to deal with some practical situations in which some face images in database are unclear or blur or the distribution of database has high bias. Denote training data as $D=\left\{\left(\mathbf{A}^{(i)},w_i\right),i=1,\dots,N\right\}$

Algorithm 1: Construct the proposed face model

- Step 1: Compute the mean image

$$\overline{\mathbf{A}}=\frac{\sum_{i=1}^{N} w_i\mathbf{A}^{(i)}}{\sum_{i=1}^{N} w_i} \quad (1)$$

- Step 2: Compute matrix

$$\mathbf{G}=\frac{\sum_{i=1}^{N} w_i\left(\mathbf{A}^{(i)}-\overline{\mathbf{A}}\right)^T\left(\mathbf{A}^{(i)}-\overline{\mathbf{A}}\right)}{\sum_{i=1}^{N} w_i} \quad (2)$$

- Step 3: Compute eigenvectors and eigenvalues of G $\left\{\mathbf{\Omega}_1,\mathbf{\Omega}_2,\dots,\mathbf{\Omega}_n\right\}$ and $\left\{\lambda_1,\lambda_2,\dots,\lambda_n\right\}$

B. *Feature extraction*

We project input image **A** on each principal component of the proposed face model.

$$\left\{\mathbf{X}_k=\mathbf{A}\mathbf{\Omega}_k,k=1,\dots,d\right\} \quad (3)$$

Finally, we have the following feature matrix $\left(\mathbf{X}_1,\mathbf{X}_2,\dots,\mathbf{X}_n\right)$.

III. Support Vector Machine

The goal of SVM classifiers is to find a hyperplane that separate the largest fraction of a labeled data set $\{(\mathbf{x}^{(i)},y^{(i)});i=1,\dots,N\}$ where $y^{(i)}$ is the associate label (-1 female or +1 male). The most important requirement, which the classifiers must have, is to maximize the distance or the margin between each class and the hyperplane (Fig 3.).

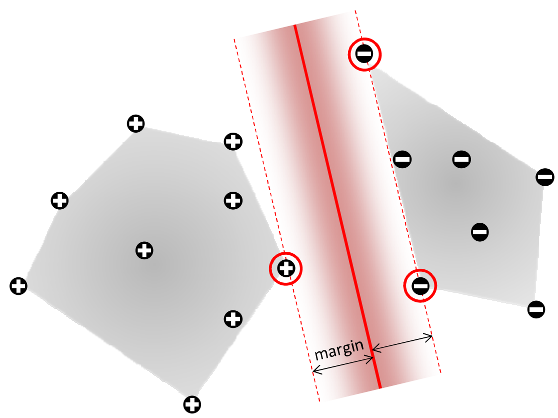

Figure 3. An SVM classifier

Clearly, most of real data is not linearly classified. To deal with difficulty, we could transform data into a high-dimensional feature space and assume that our data in this space can be linearly classified.

$$\begin{aligned} {}^{n} &\rightarrow {}^{m} \\ \mathbf{x} &\mapsto \mathbf{\Phi} \end{aligned} \quad (4)$$

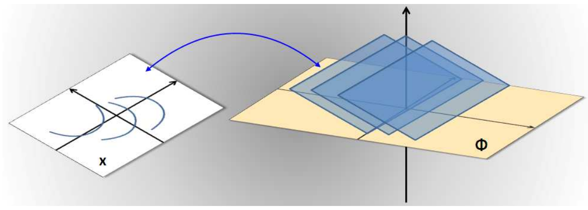

Figure 4. Input space and Feature space

In fact, determining the optimal hyperplane is a constrained optimization problem and solved using quadratic programming techniques. The discriminant hyperplane is defined by

$$y(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y^{(i)} K\left(\mathbf{x}^{(i)}, \mathbf{x}\right) + b \tag{5}$$

where $K(\mathbf{x}', \mathbf{x}'')$ is the kernel function.

Obviously, the optimal hyperplane or decision boundary mainly depends on vectors near to it. Therefore, to reduce unnecessary vectors we use the technique k-NN. It means that we will find the k nearest neighbors of a vector. If all of them belong to the class of the vector, we can drop it from the data set.

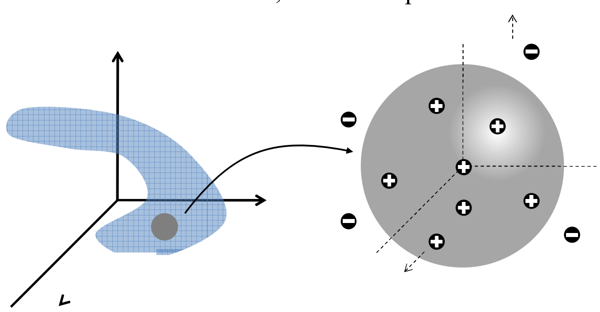

Figure 5. Data reduction

A. *Classifier Construction Phase*

Algorithm 2: Construct classifier

- Step 1: Compute H

$$H_{ij} = y^{(i)} y^{(j)} K\left(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}\right) \tag{6}$$

- Step 2: This is a quadratic optimization problem where objective function

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left(\frac{1}{2} \boldsymbol{\alpha}^T \mathbf{H} \boldsymbol{\alpha} - \sum_{i=1}^{N} \alpha_i \right)$$

$$\begin{cases} 0 \le \alpha_i \le C \\ \sum_{i=1}^{N} \alpha_i y^{(i)} = 0 \end{cases} \tag{7}$$

Use ***quadratic solvers*** to solve the optimization problem

- Step 3: Compute *b*

$$idx = \{ i \mid \alpha_i > 0 \}$$

$$N_{idx} = |idx| \tag{8}$$

$$b = \frac{1}{N_{idx}} \sum_{i \in idx} \left(y^{(i)} - \sum_{j \in idx} \alpha_j y^{(j)} K\left(\mathbf{x}^{(j)}, \mathbf{x}^{(i)}\right) \right)$$

B. *Classification Phase*

Algorithm 3: Classify

- Step 1: Compute

$$y = \operatorname{sgn}\left(\sum_{i=1}^{N} \alpha_i y^{(i)} K\left(\mathbf{x}^{(i)}, \mathbf{x}\right) + b \right) \tag{9}$$

- Step 2: Classify

$$\begin{cases} \text{if } y = +1 \text{ then } \mathbf{x} \text{ belong male} \\ \text{if } y = -1 \text{ then } \mathbf{x} \text{ belong female} \end{cases} \tag{10}$$

IV. Implementation and Experiments

We selected FERET face database to evaluate our approach. The FERET database was collected at George Mason University between August 1993 and July 1996. It contains 14,126 images stored in 1564 sets for 1199 individuals and 365 duplicate sets.

First of all, from FERET database we constructed a data set for building face models. The size of data set was 200 including 100 males and 100 females that were randomly chosen from FERET database. We used the algorithm 1 and conventional PCA to construct the models.

We designed face gender classification system including two main modules of extraction and classification. In the first module, face regions were identified and extracted from the background of the input image by using the well-known algorithm developed by Viola and Jones [11]. From the results of this step we cropped the areas of images containing faces. Feature matrices are extracted by using Formula (3). To construct the face gender classifier, we used Algorithm 2 and run it on the training data set which contained 200 individuals (100 males and 100 females) that were randomly selected from FERET database.

After training the SVM classifier, we conducted two experiments for gender classification.

Experiment 1: We applied the PCA face model to extract face features; and used two kernel functions, using Algorithm 3 to classify input features as males or females.

Polynomial $K(\mathbf{x},\mathbf{x}')=(\mathbf{x}^T\mathbf{x}'+1)^d$ with $d = 3$

Radial Basis Functions (RBF) $K(\mathbf{x},\mathbf{x}')=e^{-\left(\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{2\sigma^2}\right)}$

TABLE I. EXPERIMENT 1 RESULTS

Classifier	Error rate (in %)		
	Male	***Female***	***Overall***
RBF SVM	2.04	4.77	3.39
Polynomial SVM	4.20	5.95	4.83

Experiment 2: We applied our proposed 2DPCA face model to extract face features; and also used the same kernel functions for SVM classifiers.

TABLE II. EXPERIMENT 2 RESULTS

Classifier	Error rate (in %)		
	Male	***Female***	***Overall***
RBF SVM	2.01	4.5	3.30
Polynomial SVM	4.15	5.41	4.51

V. CONCLUSION

In summary, we have proposed a new approach for face gender classification. The first contribution of this paper is to propose a novel face model based on conventional 2DPCA. The second contribution of this paper is to combine our proposed face model with SVM. We compared our method against traditional methods. The results from our method outperformed significantly.

A future direction of research is to extend the problem to facial expression classification.

REFERENCES

[1] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on, 1991, pp. 586-591.

[2] L. Wiskott, et al., "Face recognition by elastic bunch graph matching," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, pp. 775-779, 1997.

[3] M. S. Bartlett, et al., "Face recognition by independent component analysis," Neural Networks, IEEE Transactions on, vol. 13, pp. 1450-1464, 2002.

[4] Y. Ming-Hsuan, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face recognition using kernel methods," in Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on, 2002, pp. 215-220.

[5] Y. Jian, et al., "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, pp. 131-137, 2004.

[6] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273-297, 1995.

[7] C. Huajie and W. Wei, "Pseudo-Example Based Iterative SVM Learning Approach for Gender Classification," in Intelligent Control and Automation, 2006. WCICA 2006. The Sixth World Congress on, 2006, pp. 9528-9532.

[8] B. Moghaddam and Y. Ming-Hsuan, "Gender classification with support vector machines," in Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, 2000, pp. 306-311.

[9] H. Xia, et al., "Gender Classification Based on 3D Face Geometry Features Using SVM," in CyberWorlds, 2009. CW '09. International Conference on, 2009, pp. 114-118.

[10] L. Xue-Ming and W. Yi-Ding, "Gender classification based on fuzzy SVM," in Machine Learning and Cybernetics, 2008 International Conference on, 2008, pp. 1260-1264.

[11] P. Viola and M. J. Jones, "Robust real-time face detection," International Journal of Computer Vision, vol. 57, pp. 137-154, 2004.