Research Letter

Extraction and Recognition of Nonlinear Interval-Type Features Using Symbolic KDA Algorithm with Application to Face Recognition

P. S. Hiremath[1] and C. J. Prabhakar[2]

[1] Department of Studies in Computer Science, Gulbarga University, Gulbarga,
 585-106 Karnataka, India
[2] Department of Studies in Computer Science, Kuvempu University, Shankaraghatta,
 577-451 Karnataka, India

Correspondence should be addressed to C. J. Prabhakar, psajjan@yahoo.com

We present symbolic kernel discriminant analysis (symbolic KDA) for face recognition in the framework of symbolic data analysis. Classical KDA extracts features, which are single-valued in nature to represent face images. These single-valued variables may not be able to capture variation of each feature in all the images of same subject; this leads to loss of information. The symbolic KDA algorithm extracts most discriminating nonlinear interval-type features which optimally discriminate among the classes represented in the training set. The proposed method has been successfully tested for face recognition using two databases, ORL database and Yale face database. The effectiveness of the proposed method is shown in terms of comparative performance against popular face recognition methods such as kernel Eigenface method and kernel Fisherface method. Experimental results show that symbolic KDA yields improved recognition rate.

1. INTRODUCTION

Kernel principal component analysis (KPCA) [1, 2] and kernel Fisher discriminant analysis (KFD) [3] have aroused considerable interest in the face recognation problem. KPCA was originally developed by Schölkopf et al., in 1998, while KFD was first proposed by Mika et al., in 1999 [3]. Subsequent research saw the development of series of KFD algorithms. The KFD based algorithms use the pixel intensity values in a face image as the features for face recognition. The pixel intensities that are used as features are represented by single valued variables. However, in many situations same face is captured in different orientation, lighting, expression and background, which all lead to image variations. The pixel intensities do change because of image variations. The use of single valued variables may not be able to capture the variation of feature values of the images of the same subject. In such a case, we need to consider the symbolic data analysis (SDA) [4–7], in which the interval-valued data are analyzed.

In this paper, new appearance based method is proposed in the framework of symbolic data analysis, namely, symbolic KDA for face recognition, which is a generalization of the classical KDA to symbolic objects. In the first step, we represent the face images as symbolic objects (symbolic faces) of interval type variables. Each symbolic face summarizes the variation of feature values through the different images of the same subject. It also drastically reduces the dimension of the image space without losing a significant amount of information.

In the second step, we apply symbolic KDA algorithm to extract interval type nonlinear discriminating features. According to this algorithm, in the first phase, we apply kernel function to symbolic faces, as a result a pattern in the original input space is mapped into a potentially much higher dimensional feature vector in the feature space, and then performs in the feature space to choose subspace dimension carefully. In the second phase, we extract interval type nonlinear discriminating features, which are robust to

variations due to illumination, orientation and facial expression. Finally, minimum distance classifier with symbolic dissimilarity measure [4] is employed for classification.

The remainder of this paper is organized as follows: In Section 2, the idea of constructing the symbolic faces is given. Symbolic KDA is developed in Section 3. In Section 4, the experiments are performed on the ORL and Yale face database whereby the proposed algorithm is evaluated and compared with other methods. Finally, conclusion and discussion are given in Section 5.

2. CONSTRUCTION OF SYMBOLIC FACES

Let $\Omega = \{\Gamma_1, \ldots, \Gamma_n\}$ be the collection of n face images of the database, each of size $N \times M$, which are first-order objects. Each object $\Gamma_l \in \Omega$, $l = 1, \ldots, n$, is described by a feature vector $(\widetilde{Y}_1, \ldots, \widetilde{Y}_p)$ of length $p = NM$, where each component $\widetilde{Y}_j$, $j = 1, \ldots, p$, is a single-valued variable representing the intensity values of the face image Γ_l. An image set is a collection of face images of m different subjects (face classes) and each subject has different images with varying expressions, orientations, and illuminations. Thus there are m number of second-order objects (face classes) denoted by $E = \{c_1, c_2, \ldots, c_m\}$, each consisting of different individual images, $\Gamma_l \in \Omega$, of a subject. We have assumed that images are belonging to a face class and are arranged from right-side view to left-side view. The view range of each face class is partitioned into q subface classes and each subface class contains r number of images. The feature vector of kth subface class c_i^k of ith face class c_i, where $k = 1, 2, \ldots, q$, is described by a vector of p interval variables $Y_1, \ldots, Y_p$, and is of length $p = NM$. The interval variable Y_j of kth subface class c_i^k of ith face class is described as

$$Y_j(c_i^k) = \left[\underline{x}_{ij}^k, \overline{x}_{ij}^k\right], \tag{1}$$

where $\underline{x}_{ij}^k$ and $\overline{x}_{ij}^k$ are minimum and maximum intensity values, respectively, among jth pixels of all the images of subface class c_i^k. This interval incorporates variability of jth feature inside the kth subface class c_i^k.

We denote

$$X_i^k = \left(Y_1(c_i^k), \ldots, Y_p(c_i^k)\right), \quad i = 1, \ldots, m, \ k = 1, \ldots, q. \tag{2}$$

The vector X_i^k of interval variables is recorded for kth subface class c_i^k of ith face class. This vector is called as *symbolic face* and is represented as

$$X_i^k = (\alpha_{i1}^k, \ldots, \alpha_{ip}^k), \tag{3}$$

where $\alpha_{ij}^k = Y_j(c_i^k) = [\underline{x}_{ij}^k, \overline{x}_{ij}^k]$, $j = 1, \ldots, p$, and $k = 1, \ldots, q$; $i = 1, 2, \ldots, m$.

We represent the qm symbolic faces by a matrix X of size $(p \times qm)$, consisting of column vectors X_i^k, $i = 1, \ldots, m$, $k = 1, \ldots, q$.

3. ACQUIRING NONLINEAR SUBSPACE USING SYMBOLIC KDA METHOD

Let us consider the matrix X containing qm symbolic faces pertaining to the given set Ω of images belonging to m face classes. The centers $x_{ij}^{kC} \in \Re$ of the intervals $\alpha_{ij}^k = [\underline{x}_{ij}^k, \overline{x}_{ij}^k]$ are given by

$$x_{ij}^{kC} = \frac{\overline{x}_{ij}^k + \underline{x}_{ij}^k}{2}, \tag{4}$$

where $k = 1, \ldots, q$, $i = 1, \ldots, m$, and $j = 1, \ldots, p$.

The $p \times qm$ data matrix X^C containing the centers $x_{ij}^{kC} \in \Re$ of the intervals for qm symbolic faces. The p-dimensional vectors $X_i^{kC} = (x_{i1}^{kC}, \ldots, x_{ip}^{kC})$, $\underline{X}_i^k = (\underline{x}_{i1}^k, \ldots, \underline{x}_{ip}^k)$, and $\overline{X}_i^k = (\overline{x}_{i1}^k, \ldots, \overline{x}_{ip}^k)$ represent the centers, lower bounds, and upper bounds of the qm symbolic faces X_i^k, respectively.

Let $\Phi : \Re^p \rightarrow F$ be a nonlinear mapping between the input space and the feature space.

The nonlinear mapping, Φ, usually defines a kernel function. Let $K \in \Re^{qm \times qm}$ define a kernel matrix by means of dot product in the feature space:

$$K_{ij} = \left(\Phi(X_i) \cdot \Phi(X_j)\right). \tag{5}$$

In general, the Fisher criterion function in the feature space F can be defined as

$$J(V) = \frac{V^T S_b^\Phi V}{V^T S_w^\Phi V}, \tag{6}$$

where V is a discriminant vector, S_b^Φ and S_w^Φ are the between-class scatter matrix and the within-class scatter matrix, respectively. The between-class scatter matrix and the within-class scatter matrix in the feature space F are defined below:

$$S_b^\Phi = \frac{1}{m} \sum_{i=1}^m q_i (m_i^\Phi - m_0^\Phi)(m_i^\Phi - m_0^\Phi)^T,$$

$$S_w^\Phi = \frac{1}{qm} \sum_{i=1}^m \sum_{k=1}^q (\Phi(X_i^{kC}) - m_i^\Phi)(\Phi(X_i^{kC}) - m_i^\Phi)^T, \tag{7}$$

where X_i^{kC} denotes the kth symbolic face of ith face class, q_i is the number of training symbolic faces in face class i, m_i^Φ is the mean of the mapped symbolic faces in face class i, and m_0^Φ is the mean across all mapped qm symbolic faces. From the above definitions, we have $S_t^\Phi = S_b^\Phi + S_w^\Phi$. The discriminant vectors with respect to the Fisher criterion are actually the eigenvectors of the generalized equation $S_b^\Phi V = \lambda S_t^\Phi V$. According to the theory of the reproducing kernel, V will be an expansion of all symbolic faces in the feature space. That is, there exist coefficients b_L ($L = 1, 2, \ldots, qm$) such that

$$V = \sum_{L=1}^{qm} b_L \Phi(X_i^{kC}) = HA, \tag{8}$$

where $H = (\Phi(X_1^{1C}), \ldots, \Phi(X_1^{qC}), \ldots, \Phi(X_m^{1C}), \ldots, \Phi(X_m^{qC}))$ and $A = (b_1 b_2 \cdots b_{qm})$.

Substituting (8) into (6), we can obtain the following equation:

$$J(A) = \frac{A^T K W K A}{A^T K K A}, \qquad (9)$$

where K is a kernel matrix, $W = \text{diag}(W_1 \cdots W_m)$, W_i is a $q_i \times q_i$ matrix whose elements are $1/q_i$. From the definition of W, it is easy to verify that W is a $qm \times qm$ block matrix. In fact, it is often necessary to find s discriminant vectors, denoted by $a_1, a_2, \ldots, a_s$, to extract features. Let $V = [a_1, a_2, \ldots, a_s]$. The matrix V should satisfy the following condition:

$$V = \arg\max_V \left(\frac{|V^T \widetilde{S}_b V|}{|V^T \widetilde{S}_t V|} \right), \qquad (10)$$

where $\widetilde{S}_b = KWK$ and $\widetilde{S}_t = KK$.

Since, each symbolic face X_i^k is located between the lower bound symbolic face $\underline{X}_i^k$ and upper bound symbolic face $\overline{X}_i^k$, it is possible to find most discriminating nonlinear interval-type features $[\underline{B}_i^k, \overline{B}_i^k]$.

The lower bound features of each symbolic face X_i^k is given by

$$\underline{B}_i^k = V_l^T \Phi(\underline{X}_i^k), \quad l = 1, 2, \ldots, s, \qquad (11)$$

where

$$\Phi(\underline{X}_i^k) = \left[\left(\Phi(X_1^{1^C}) \cdot \Phi(\underline{X}_i^k), \ldots, \Phi(X_1^{q^C}) \cdot \Phi(\underline{X}_i^k) \right), \right.$$
$$\left. \ldots, \left(\Phi(X_m^{1^C}) \cdot \Phi(\underline{X}_i^k), \ldots, \Phi(X_m^{q^C}) \cdot \Phi(\underline{X}_i^k) \right) \right]. \qquad (12)$$

Similarly, the upper bound features of each symbolic face X_i^k is given by

$$\overline{B}_i^k = V_l^T \Phi(\overline{X}_i^k), \quad l = 1, 2, \ldots, s. \qquad (13)$$

Let $c_{\text{test}} = [\Gamma_1, \Gamma_2, \ldots, \Gamma_l]$ be the test face class that contains face images of same subject with different expression, lighting condition and orientation. The test symbolic face X_{test} is constructed for test face class c_{test} as explained in Section 2. The lower bound test symbolic face of test symbolic face X_{test} is described as $\underline{X}_{\text{test}} = (\underline{x}_1^{\text{test}}, \underline{x}_2^{\text{test}}, \ldots, \underline{x}_p^{\text{test}})$. Similarly, the upper bound test symbolic face is described as $\overline{X}_{\text{test}} = (\overline{x}_1^{\text{test}}, \overline{x}_2^{\text{test}}, \ldots, \overline{x}_p^{\text{test}})$.

The interval-type features $[\underline{B}^{\text{test}}, \overline{B}^{\text{test}}]$ of test symbolic face X_{test} are computed as:

$$\underline{B}^{\text{test}} = V_l^T \Phi(\underline{X}_{\text{test}}),$$
$$\overline{B}^{\text{test}} = V_l^T \Phi(\overline{X}_{\text{test}}), \qquad (14)$$

where $l = 1, 2, \ldots, s$.

4. EXPERIMENTAL RESULTS

The proposed symbolic KDA method is experimented with the face images of the ORL and Yale databases. The effectiveness of proposed method is shown in terms of comparative performance against two face recognition methods. In particular, we compared our algorithm with kernel Eigenface [8] and kernel Fisherface [9] method.

FIGURE 1: Some typical images of one subject of ORL database.

TABLE 1: Optimal parameters corresponding to each method with respect to two different kernels.

Method	Polynomial kernel		Gaussian kernel	
	Order	Subspace dimension	Width	Subspace dimension
Kernel Eigenfaces	1	44	7	47
Kernel fisherfaces	3	35	5	44
Symbolic KDA	3	18	4	24

4.1. Experiments using ORL database

We assess the feasibility and performance of the proposed symbolic KDA on the face recognition task using ORL database. The ORL face database is composed of 400 images with ten different facial views that represent various expressions and orientations for each of the 40 distinct subjects as shown in Figure 1. We have arranged images of each face class from right side view to left side view. All the 400 images from the ORL database are used to evaluate the face recognition performance of proposed method. Six images are randomly chosen from the ten images available for each subject for training, while the remaining images are used to construct the test symbolic face for each trial. We have conducted the experiments using two kernel functions namely, polynomial kernel and Gaussian kernel.

Our goal is to find appropriate kernel function and corresponding optimal kernel parameters (i.e., the order of the polynomial kernel and the width of the Gaussian kernel) for our proposed method. The experimental results shows that the order of the polynomial kernel should be three and the width of Gaussian kernel should be four for proposed symbolic KDA with respect to a minimum distance classifier.

After determining the optimal kernel parameters, we set out to select the dimension of discriminant subspace with respect to two different kernels. Table 1 shows optimal subspace for each method. The optimal parameters for each method with respect to different kernels. Besides, we find that symbolic KDA features seem more effective than features of other methods.

After selection of optimal parameters and optimal subspace for each method with respect to different kernels, all three methods are reevaluated using same set of training and testing samples. The average recognition rates for the best case are shown in Table 2. The best performance of the symbolic KDA method is better than the best performance of the kernel Eigenface and kernel Fisherface method. We note that the symbolic KDA method outperforms kernel Eigenface method and kernel Fisherface in the sense of using small number of features.

TABLE 2: Comparison of symbolic KDA method using optimal parameters.

	Kernel Eigenface	Symbolic KDA	Kernel Fisherface
Polynomial kernel	55.61	91.86	84.95
Gaussian kernel	58.05	89.39	81.57

FIGURE 3: Some typical images of one subject of Yale face database.

TABLE 3: Comparison of classification performance using Yale face database.

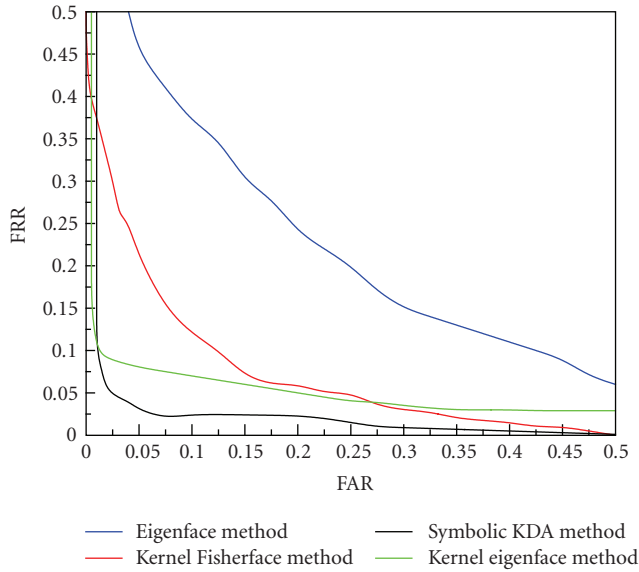| Methods	kernel Fisherfaces	kernel Eigenfaces	symbolic KDA
Recognition rates (%)	87.15	83.28	89.00
Training time (seconds)	48.346	68.751	38.324
Feature dimension	42	54	15

FIGURE 2: The ROC performance of proposed symbolic KDA Method, Kernel Eigenface Method, Kernel Fisherface Method and Eigenface Method.

In order to examine, whether symbolic KDA is statistically significant and better than other methods in terms of its recognition rate. We evaluate the experimental results presented in Table 2 using McNemar's significance test. McNemar's test is essentially a null hypothesis statistical test based on a Bernoulli model. If the resulting p-value is below the desired significance level, the null hypothesis is rejected and the performance difference between two algorithms is considered to be statistically significant. By this test, we find that symbolic KDA statistically significant and outperforms kernel Eigenface and kernel Fisherface methods at a significance level of $p = 1.012 \times 10^{-6}$.

The receiver operating characteristic (ROC) curve in Figure 2 reports results for a verification scenario. The equal-error rate (EER) is the ROC point at which the false-accept rate is equal to the false-reject rate. The EER for our approach goes from approximately 0.12 to approximately 0.23. The EER for the other methods shows much greater performance degradation.

4.2. Experiments on the Yale face database

The experiments were conducted using Yale database to evaluate the excellence of the symbolic KDA for frontal face recognition under different nondark backgrounds. The Yale face database consists of a total of 165 images obtained from 15 different people, with 11 images from each person. Figure 3 shows some typical images of one subject of Yale face database. We preprocessed these images by aligning and scaling them so that the distances between the eyes were the same for all images and also ensuring that the eyes occurred in the same co-ordinates of the image. The resulting image was then cropped. The images were not manually arranged as done in previous set of experiments using ORL database (Section 4.1). In our experiments, 9 images were randomly chosen from each class for training, while the remaining two images were used to construct test symbolic face for each trial.

The experiments were conducted using two different kernels, namely, polynomial kernel and Gaussian kernel. The order of the polynomial kernel should be 2 and the width of Gaussian kernel should be four for proposed symbolic KDA with respect to a minimum distance classifier.

After finding optimal kernel parameter (degree 2) for the symbolic KDA method, the experiments were conducted to find optimal subspace for proposed symbolic KDA, kernel Fisherface, and kernel Eigenface method. The recognition rates, training time, and optimal subspace dimension for Kernel Fisherface, Kernel Eigenface, and symbolic KDA are listed in Table 3. From Table 3, the symbolic KDA method with polynomial degree two has recognition rate 89.00% using only 15 features, where as Kernel Fisherface method requires 42 features to achieve 87.15% recognition rate. This is due to the fact that first few eigenvectors of symbolic PCA account for highest variance of training samples and these few eigenvectors are enough to represent image for recognition purposes. Hence, improved recognition results can be achieved at less computational cost by using symbolic KDA, by virtue of its low dimensionality. The experimental results obtained using Yale face database shows that the proposed symbolic KDA performs well on images with non dark backgrounds.

5. CONCLUSIONS

In this paper, we propose a novel symbolic KDA method for face recognition. Symbolic data representation of face images using interval-type features are desirable facial features to cope up with the variations due to illumination, orientation, and facial expression changes. The feasibility of the symbolic KDA has been tested successfully on frontal face images of ORL and Yale databases. Experimental results show that symbolic KDA method with polynomial kernel leads to improved recognition rate at reduced computational cost.

The proposed symbolic KDA has many advantages compared to other popular appearance-based methods. The drawback of other appearance-based methods is that in order to recognize a face seen from a particular pose and under a particular illumination, the face must have been previously seen under the same conditions. The symbolic KDA overcomes this limitation by representing the faces by interval-type features so that even the faces seen previously in different poses, orientations, and illuminations are recognized. Another important merit is that we can use more than one probe image with inherent variability of a face for face recognition, this yields improved recognition rate. This is clearly evident from the experimental results. We observe from the experimental results that the proposed symbolic KDA method yields improved recognition rate in terms of time and feature reduction compare to other kernel-based methods.

The main drawback of our proposed symbolic KDA method is that pose variation is limited up to 20 degree orientation and the performance of proposed method decreases on face images with pose variation greater than 20 degree orientation. The proposed method did not achieve 100% accuracy, this is due to the fact that while constructing the symbolic faces, there may be chance of misalignment of coordinates of eyes, mouth, and nose because of different facial expressions in training images. It can be observed in experimental results obtained using Yale face database, which contains face images with different facial expressions. Moreover, the performance of the proposed symbolic KDA method decreases on images with more variation in facial expressions of Yale face database compared to performance on images with less variation in facial expressions of ORL face database.

ACKNOWLEDGMENT

REFERENCES

[1] K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 181–201, 2001.

[2] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[3] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K.-R. Müller, "Fisher discriminant analysis with kernels," in *Proceedings of the 9th IEEE Workshop on Neural Networks for Signal Processing (NNSP '99)*, pp. 41–48, Madison, Wis, USA, August 1999.

[4] H.-H. Bock and E. Diday, Eds., *Analysis of Symbolic Data*, Springer, Heidelberg, Germany, 2000.

[5] P. S. Hiremath and C. J. Prabhakar, "Symbolic factorial discriminant analysis for face recognition under variable lighting," in *Proceedings of the 3rd International Conference on Visual Information Engineering (VIE '06)*, pp. 150–154, IET, Bangalore, India, September 2006.

[6] P. S Hiremath and C. J. Prabhakar, "Acquiring non linear subspace for face recognition using symbolic kernel PCA method," *Journal of Symbolic Data Analysis*, vol. 4, no. 1, pp. 15–26, 2006.

[7] P. S. Hiremath, A. Dhanti, and C. J. Prabhakar, "Modeling uncertainty in representation of facial features for face recognition," in *Face Recognition*, K. Delac and M. Grgic, Eds., chapter 10, pp. 183–218, ITech, Vienna, Austria, 2007.

[8] M.-H. Yang, N. Ahuja, and D. Kriegman, "Face recognition using kernel eigenfaces," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '00)*, vol. 1, pp. 37–40, Vancouver, BC, Canada, September 2000.

[9] M.-H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods," in *Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition (FGR '02)*, pp. 215–220, Washington, DC, USA, May 2002.