

RESEARCH

Open Access

Facial expression recognition using local binary patterns and discriminant kernel locally linear embedding

Xiaoming Zhao¹ and Shiqing Zhang^{2*}

Abstract

Given the nonlinear manifold structure of facial images, a new kernel-based supervised manifold learning algorithm based on locally linear embedding (LLE), called discriminant kernel locally linear embedding (DKLLE), is proposed for facial expression recognition. The proposed DKLLE aims to nonlinearly extract the discriminant information by maximizing the interclass scatter while minimizing the intraclass scatter in a reproducing kernel Hilbert space. DKLLE is compared with LLE, supervised locally linear embedding (SLLE), principal component analysis (PCA), linear discriminant analysis (LDA), kernel principal component analysis (KPCA), and kernel linear discriminant analysis (KLDA). Experimental results on two benchmarking facial expression databases, i.e., the JAFFE database and the Cohn-Kanade database, demonstrate the effectiveness and promising performance of DKLLE.

Keywords: manifold learning, locally linear embedding, facial expression recognition

Introduction

Affective computing, which is currently an active research area, aims at building the machines that recognize, express, model, communicate and respond to a user's emotion information [1]. Within this field, recognizing human emotion from facial images, i.e., facial expression recognition, is increasingly attracting attention and has become an important issue, since facial expression provides the most natural and immediate indication about a person's emotions and intentions. Over the last decade, the importance of automatic facial expression recognition has increased significantly due to its applications to human-computer interaction (HCI), human emotion analysis, interactive video, indexing and retrieval of image, etc.

An automatic facial expression recognition system generally comprises of three crucial steps [2]: face acquisition, facial feature extraction, and facial expression classification. Face acquisition is a preprocessing stage to detect or locate the face regions in the input images or sequences. One of the most widely used face detector is the real-time face detection algorithm developed by Viola and Jones [3], in

which a cascade of classifiers is employed with Harr-wavelet features. Once a face is detected in the images, the corresponding face regions are usually normalized to have the same eye distance and the same gray level. Facial feature extraction attempts to find the most appropriate representation of facial images for recognition. There are mainly two approaches: geometric features-based systems and appearance features-based systems. In the geometric features-based systems, the shape and locations of major facial components such as mouth, nose, eyes, and brows, are detected in the images. Nevertheless, the geometric features-based systems require the accurate and reliable facial feature detection, which is difficult to realize in real-time applications. In the appearance features-based systems, the appearance changes (skin texture) of the facial images, including wrinkles, bulges, and furrows, are presented. Image filters, such as principal component analysis (PCA) [4], linear discriminant analysis (LDA) [5], regularized discriminant analysis (RDA) [6] and Gabor wavelet analysis [7,8], can be applied to either the whole-face or specific face regions to extract the facial appearance changes. It's worth pointing out that it is computationally expensive to convolve facial images with a set of Gabor filters to extract multi-scale and multi-orientation coefficients. Moreover, in practice the dimensionality of Gabor features is so high

* Correspondence: tzcqsq@163.com

²School of Physics and Electronic Engineering, Taizhou University, Taizhou 318000, P.R. China

Full list of author information is available at the end of the article

that the computation and memory requirements are very large. In recent years, an effective face descriptor called local binary patterns (LBP) [9], originally proposed for texture analysis [10], have attracted extensive interest for facial expression representation. One of the most important properties of LBP is its tolerance against illumination changes and its computational simplicity. So far, LBP has been successfully applied as a local feature extraction method in facial expression recognition [11-13]. In the last step of an automatic facial expression recognition system, i. e., facial expression classification, a classifier is employed to identify different expressions based on the extracted facial features. The representative classifiers used for facial expression recognition are neural networks [14], the nearest neighbor (1-NN) [15] or k-nearest neighbor (KNN) classifier [16], and support vector machines (SVM) [17], etc.

In recent years, it has been proved that facial images of a person with varying expressions can be represented as a low-dimensional nonlinear manifold embedded in a high-dimensional image space [18-20]. Given the nonlinear manifold structure of facial expression images, two representative manifold learning (also called nonlinear dimensionality reduction) methods, i.e., locally linear embedding (LLE) [21] and isometric feature mapping (Isomap) [22], have been used to project the high-dimensional facial expression images into a low-dimensional embedded subspace in which facial expressions can be easily distinguished from each other [18-20,23,24]. However, LLE and Isomap fail to perform well on facial expression recognition tasks due to their unsupervised ways of failing to extract the discriminant information.

To overcome the limitations of unsupervised manifold learning methods for supervised pattern recognition, some supervised manifold learning algorithms have been recently proposed by means of a supervised distance measure, such as supervised locally linear embedding (SLLE) [25] using the linear supervised distance, probability-based LLE using a probability-based distance [26], locally linear discriminant embedding using a vector translation and distance rescaling model [27], and so forth. Among them, SLLE has become one of the most promising supervised manifold learning techniques due to its simple implementation, and successfully applied for facial expression recognition [28]. However, SLLE still has two shortcomings. Firstly, due to the used linear supervised distance, the interclass dissimilarity in SLLE keeps increasing in parallel while the intraclass dissimilarity is increased. However, an ideal classification mechanism should maximize the interclass dissimilarity while minimizing the intraclass dissimilarity. In this sense, this kind of linear supervised distance in SLLE is not a good property for classification since it will go to a great extent to decrease the discriminating power of the low-dimensional embedded data representations produced with SLLE. Secondly, as a non-kernel method,

SLLE cannot explore the higher-order information of input data as SLLE cannot employ the characteristic of a kernel-based learning, i.e., a nonlinear kernel mapping. To tackle the above-mentioned problems of SLLE, in this article a new kernel-based supervised manifold learning algorithm based on LLE, called discriminant kernel locally linear embedding (DKLLE), is proposed and applied for facial expression recognition. On one hand, with a nonlinear supervised distance measure, DKLLE considers both the intraclass scatter information and the interclass scatter information in a reproducing kernel Hilbert space (RKHS), and emphasizes the discriminant information. On the other hand, with kernel techniques DKLLE extracts the nonlinear feature information when mapping input data into some high dimensional feature space. In order to evaluate the performance of DKLLE on facial expression recognition, we adopt the LBP features as facial representations and then employ DKLLE to produce the low-dimensional discriminant embedded data representations from the extracted LBP features with striking performance improvement on facial expression recognition tasks. The facial expression recognition experiments are performed on two benchmarking facial expression databases, i.e., the JAFFE database [15] and the Cohn-Kanade database [29].

The remainder of this article is organized as follows: in Section 2, LBP is introduced briefly. In Section 3, LLE and SLLE are reviewed briefly. The proposed DKLLE algorithm is presented in detail in Section 4. In Section 5, experiments and results are given. Finally, the conclusions are summarized in Section 6.

Local binary patterns

The original LBP operator [10] labels the pixels of an image by thresholding a 3×3 neighborhood of each pixel with the center value and considering the results as a binary code. The LBP code of the center pixel in the neighborhood is obtained by converting the binary code into a decimal one. Figure 1 gives an illustration for the basic LBP operator. Based on the operator, each pixel of an image is labeled with an LBP code. The 256-bin histogram of the labels contains the density of each label and can be used as a texture descriptor of the considered region.

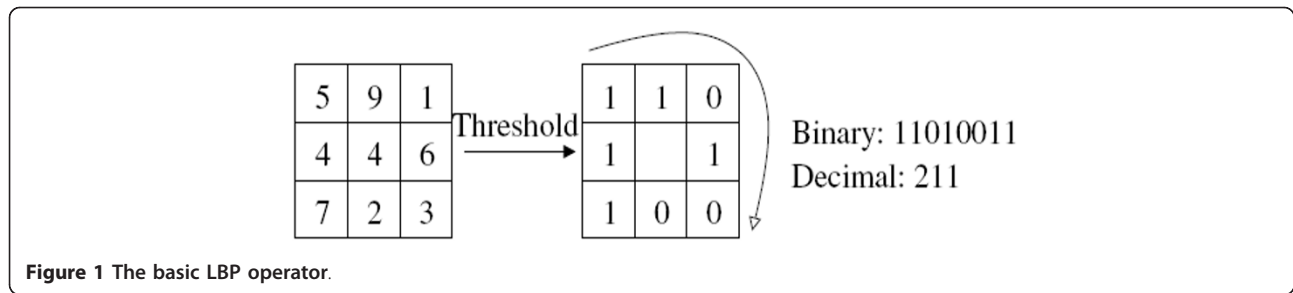
The procedure of extracting LBP features for facial representations is implemented as follows:

First, a face image is divided into several non-overlapping blocks. Second, LBP histograms are computed for each block. Finally, the block LBP histograms are concatenated into a single vector. As a result, the face image is represented by the LBP code.

LLE and SLLE

LLE

Given the input data point $x_i \in R^D$ and the output data point $y_i \in R^d$ ($i = 1, 2, 3, \dots, N$), the standard LLE [21]



consists of three steps:

Step 1: Find the number of nearest neighbors for each x_i based on the Euclidean distance.

Step 2: Compute the reconstruction weights by minimizing the reconstruction error.

Let x_i and x_j be neighbors, the reconstruction error is measured by the following cost function:

$$\varepsilon(W) = \sum_{i=1}^N \left\| x_i - \sum_{j=1}^N W_{ij} x_j \right\|^2 \quad (1)$$

subject to two constraints: $\sum_{j=1}^N W_{ij} = 1$ and $W_{ij} = 0$, if x_i and x_j are not neighbors.

Step 3: Compute the low-dimensional embedding.

The low-dimensional embedding is found through the following minimization:

$$\phi(Y) = \sum_{i=1}^N \left\| y_i - \sum_{j=1}^N W_{ij} y_j \right\|^2 \quad (2)$$

subject to two constraints: $\sum_{i=1}^N y_i = 0$ and $\frac{1}{N} \sum_{i=1}^N y_i^T y_i = I$, where I is the $d \times d$ identity matrix. To find the matrix Y under these constraints, a new matrix M is constructed based on the matrix W : $M = (I - W)^T (I - W)$. The d eigenvectors which correspond to the d smallest non-zero eigenvalues of M yield the final embedding Y .

SLLE

To complement the original LLE, SLLE [25] aims to find a mapping separating within-class structure from a between-class structure. One way to do this is to add the distance between samples x_i and x_j in different classes to modify the first step of the original LLE, while leaving the other two steps unchanged. This can be achieved by artificially increasing the pre-calculated

Euclidean distance (abbreviated as Δ) between samples belonging to different classes, but leaving them unchanged if samples are from the same class:

$$\Delta' = \Delta + \alpha \max(\Delta) \Lambda_{ij}, \alpha \in [0, 1] \quad (3)$$

where Δ is the distance matrix without considering the class label information, and Δ' is the distance integrating with the class label information. If x_i and x_j belong to the different classes, then $\Lambda_{ij} = 1$ and $\Lambda_{ij} = 0$ otherwise. In this formulation, the constant factor α ($0 \leq \alpha \leq 1$) controls the amount to which the class information should be incorporated. At one extreme, when $\alpha = 0$, we get the unsupervised LLE. At the other extreme, when $\alpha = 1$, we get the fully supervised LLE (1-SLLE). As α varies between 0 and 1, a partially supervised LLE (α -SLLE) is obtained. From Eq. (3), it can be observed that when the intraclass dissimilarity (i.e., $\Delta' = \Delta$, when $\Lambda_{ij} = 0$) is linearly increased, the interclass dissimilarity (i.e., $\Delta' = \Delta + \alpha \max(\Delta)$, when $\Lambda_{ij} = 1$) keeps increasing in parallel, since $\alpha \max(\Delta)$ is a constant. Therefore, the used supervised distance measure in SLLE is linear.

The proposed DKLLE

A discriminant and kernel variant of LLE is developed by designing a nonlinear supervised distance measure and minimizing the reconstruction error in a RKHS, which gives rise to DKLLE.

Given the input data point (x_i, L_i) , where $x_i \in R^D$ and L_i is the class label of x_i , the output data point is $y_i \in R^d$ ($i = 1, 2, 3, \dots, N$). The detailed steps of DKLLE are presented as follows:

Step 1: Perform the kernel mapping for each data point x_i .

A nonlinear mapping ϕ is defined as: $\phi: R^D \rightarrow \mathcal{F}, x \mapsto \phi(x)$

The input data point x_i is mapped with the nonlinear mapping ϕ , into some potentially high-dimensional feature space \mathcal{F} . Then, an inner product $\langle \cdot, \cdot \rangle$ can be defined

on \mathcal{F} for a chosen ϕ , which makes a so-called RHKS. In a RHKS, a kernel function $\kappa(x_p, x_j)$ can be defined as:

$$\kappa(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle = \varphi(x_i)^T \varphi(x_j) \quad (4)$$

where κ is called a kernel.

Step 2: Find the nearest neighbors for each $\phi(x_i)$ by using a nonlinear supervised kernel distance.

The kernel Euclidean distance measure [30] for two data points x_i and x_j induced by a kernel κ can be defined as:

$$Dist(x_i, x_j) = \sqrt{\langle \varphi(x_i) - \varphi(x_j), \varphi(x_i) - \varphi(x_j) \rangle} = \sqrt{\kappa(x_i, x_i) - 2\kappa(x_i, x_j) + \kappa(x_j, x_j)} \quad (5)$$

Let $Dist$ denotes the kernel distance matrix for all the input data points, i.e., $Dist = Dist(x_i, x_j)$, $i, j = 1, 2, \dots, N$. To preserve the intraclass neighboring geometry, while maximizing the interclass scatter, a nonlinear supervised kernel distance measure $KDist$ in a RHKS can be defined as:

$$KDist = \begin{cases} \sqrt{1 - e^{-\frac{Dist^2}{\beta}}} & L_i = L_j \\ \sqrt{e^{-\frac{Dist^2}{\beta}} - \alpha} & L_i \neq L_j \end{cases} \quad (6)$$

where $KDist$ is the supervised kernel distance matrix with the class label information, while $Dist$ is the kernel Euclidean distance matrix without the class label information. α is a constant factor ($0 \leq \alpha \leq 1$) and gives a certain chance for the data points in different classes to be more similar so that the dissimilarity in different classes may be smaller than that in the same class. β is used to prevent the supervised kernel distance matrix $KDist$ from increasing too fast when the kernel Euclidean distance matrix $Dist$ is relatively large, since $Dist$ is in the exponent. Hence, the value of β should depend on the "density" of data sets and it is usually feasible to set β to be the average kernel Euclidean distance between all pairs of data points.

As shown in Eq. (6), we can make two observations. First, both the interclass dissimilarity and the intraclass dissimilarity in $KDist$, is monotone increasing with respect to the kernel Euclidean distance. This ensures that the main geometric structure of the original data sets can be preserved well in the process of dimensionality reduction. Second, the interclass dissimilarity in $KDist$ can be always definitely larger than the intraclass dissimilarity, conferring a high discriminating power of DKLLE's low-dimensional embedded data representations. This is a good property for classification.

Step 3: Measure the reconstruction error in a RHKS. The reconstruction error is measured by the following cost function:

$$\varepsilon(W) = \sum_{i=1}^N \left\| \varphi(x_i) - \sum_{j=1}^k W_{ij} \varphi(x_{j,i}) \right\|^2 \quad (7)$$

where k is the number of nearest neighbors. Given the constraint: $\sum_{j=1}^k W_{ij} = 1$, the reconstruction error can be rewritten as follows:

$$\varepsilon(W) = \sum_{i=1}^N \left\| \varphi(x_i) - \sum_{j=1}^k W_{ij} \varphi(x_{j,i}) \right\|^2 = \sum_{i=1}^N \left\| \sum_{j=1}^k W_{ij} (\varphi(x_i) - \varphi(x_{j,i})) \right\|^2 = \sum_{i=1}^N \varepsilon(W_i) \quad (8)$$

Let $P_i = [\phi(x_i) - \phi(x_{1,i}), \phi(x_i) - \phi(x_{2,i}), \dots, \phi(x_i) - \phi(x_{k,i})]$, then

$$\varepsilon(W_i) = \left\| \sum_{j=1}^k W_{ij} (\varphi(x_i) - \varphi(x_{j,i})) \right\|^2 = \|PW_i\|^2 = W_i^T P^T P W_i = W_i^T K W_i \quad (9)$$

where $K = P^T P$ is a positive semi-definite kernel matrix. To compute the optimal weight W_i , the following Lagrange function is formulated with the constraint $W_i^T \mathbf{1} = 1$. ($\mathbf{1} = (1, 1, \dots, 1)^T$)

$$\begin{aligned} L(W_i, \lambda) &= W_i^T K W_i - \lambda (W_i^T \mathbf{1} - 1) \\ \Rightarrow W_i &= \frac{K^{-1} \mathbf{1}}{\mathbf{1}^T K^{-1} \mathbf{1}} \end{aligned} \quad (10)$$

Performing eigen-decomposition, let $K = U^T \Lambda U$, then

$$W_i = \frac{U^T \Lambda^{-1} U \mathbf{1}}{\mathbf{1}^T U^T \Lambda^{-1} U \mathbf{1}} \quad (11)$$

Therefore, the reconstruction weights can be computed by the kernel matrix's eigenvalues and eigenvectors.

Step 4: Compute the final embedding.

As LLE done, the following embedding cost function is minimized.

$$\phi(Y) = \sum_{i=1}^N \left\| y_i - \sum_{j=1}^N W_{ij} y_j \right\|^2 = \text{tr}(Y^T M Y) \quad (12)$$

where $M = (I - W)^T (I - W)$, subject to two constraints: $\sum_{i=1}^N y_i = 0$ and $\frac{1}{N} \sum_{i=1}^N y_i^T y_i = I$. The final embedding Y comprises d eigenvectors corresponding to d smallest non-zero eigenvalues of M

Experiments and results

To verify the effectiveness of the proposed DKLLE, we use two benchmarking facial expression expression databases, i.e.,

the JAFFE database [15] and the Cohn-Kanade Database [29], for facial expression recognition experiments. Each database contains seven emotions: anger, joy, sad, neutral, surprise, disgust, and fear. The performance of DKLLE is compared with LLE, SLLE, PCA, LDA, kernel principal component analysis (KPCA) [31], and kernel linear discriminant analysis (KLDA) [32]. The typical Gaussian kernel $\kappa(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$ is used for KPCA, KLDA, and DKLLE, and the parameter σ is empirically set to 1 for its satisfying performance. The number of nearest neighbors for LLE, SLLE, and DKLLE is fixed with an adaptive neighbor selection technique [33]. To cope with the embeddings of the new samples, the out-of-sample extensions of LLE and SLLE are developed by an existed linear generalization technique [34], in which a linear relation is built between the high and low-dimensional spaces and then the adaptation to a new sample can be done by updating the weight matrix W . As a kernel method, the proposed DKLLE can directly project the new samples into a low-dimensional space by using a kernel trick as in KPCA. For simplicity, the nearest neighbor (1-NN) classifier with the Euclidean metric is used for facial expression classification. A 10-fold cross validation scheme is employed in 7-class facial expression recognition experiments, and the average recognition results are reported.

Due to the computation complexity constraint, the reduced dimension is confined to the range [2, 100] with an interval of 5. An exception is that in the low range [2,10] we present the recognition results of each reduced dimension with a small interval of 1, since the reduced dimension of LDA and KLDA is at most $c-1$, where c is the number of facial expression classes. In each reduced dimension, the constant α ($0 \leq \alpha \leq 1$) for SLLE and DKLLE can be optimized using a simple exhaustive search within a scope ($\alpha = 0, 0.1, 0.2, \dots, 1$).

Preprocessing

As done in [11,12], on the JAFFE database and the Cohn-Kanade Database, the eye distance of facial images was normalized to a fixed distance of 55 pixels once the centers of two eyes were located. Generally, it is observed that the width of a face is roughly two times of the distance, and the height is roughly three times. Therefore, based on the normalized value of the eye distance, a resized image of 110×150 pixels was cropped from the original images. To locate the centers of two eyes, automatic face registration was performed by using the robust real-time face detector developed by Viola and Jones [3]. From the results of automatic face detection including face location, face width, and face height, two square bounding boxes for left eye and right eye were automatically constructed by using the geometry of

a typical up-right face which has been widely used to find a proper spatial arrangement of facial features [35]. Then, the center locations of two eyes could be automatically worked out in terms of the centers of two square bounding boxes for left eye and right eye. No further alignment of facial features such as alignment of mouth was performed. Additionally, there was no attempt made to remove illumination changes due to LBP's gray-scale invariance.

When the facial images of 110×150 pixels, including mouth, eyes, brows, and noses, were cropped from the original images, the LBP operator was applied to each cropped image and extracted the LBP features. As suggested in [10-12], we selected the 59-bin LBP operator, and divided the 110×150 pixels facial images into $42 (6 \times 7)$ blocks, and finally extracted the LBP features represented by the length of 2478 (59×42).

Experiments on the JAFFE database

The JAFFE database [15] contains 213 images of female facial expressions. Each image has a resolution of 256×256 pixels. A few examples of facial expression images from the JAFFE database are shown in Figure 2. The number of images corresponding to each of the seven categories of expressions is roughly the same. The recognition results obtained by each method at different reduced dimensions are given in Figure 3. The best results and the standard deviations (std) for different methods with the corresponding reduced dimension are listed in Table 1.

From the results in Figure 3 and Table 1, we can see that DKLLE achieves the highest accuracy of 84.06% at 40 reduced dimension, outperforming the other methods. More crucially, DKLLE makes about 9% improvement over LLE and about 6% improvement over SLLE. This demonstrates that DKLLE is able to extract the most discriminative low-dimensional embedded data representations for facial expression recognition. Note that it's difficult to perform directly a comparison with all the previously reported work on the JAFFE database due to the different experimental settings. Nevertheless, in our work with LBP-based 1-NN the reported accuracy of 84.06% is still very encouraging compared with the previously published work [12] similar to our experimental settings. In [12], after extracting the most discriminative LBP (called boosted-LBP) features, they used SVM and separately obtained 7-class facial expression recognition accuracy of 79.8, 79.8, and 81.0% with linear, polynomial, and radial basis function (RBF) kernels. It's worth pointing out that in this work for simplicity we did not use the boosted-LBP features and SVM. To further compare the performance of DKLLE with the work in [12], we will explore the performance of the

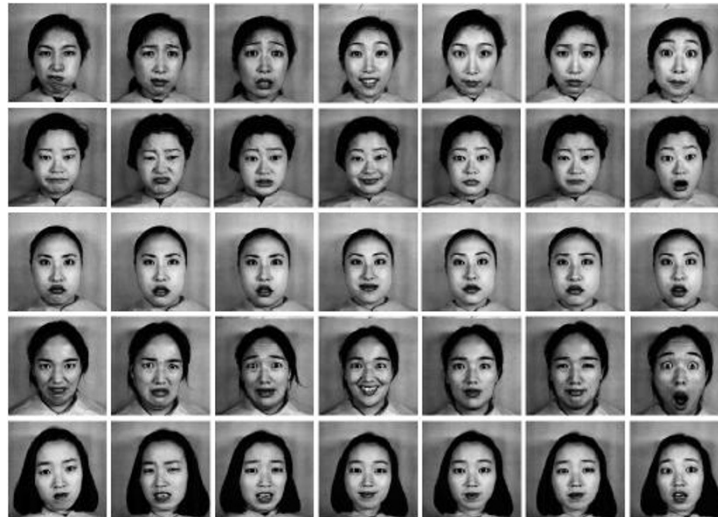


Figure 2 Examples of facial expression images from the JAFFE database.

boosted-LBP features and SVM integrating with DKLLE in our future work.

When DKLLE performs best at 40 reduced dimension, the corresponding confusion matrix of 7-class facial expression recognition results is presented in Table 2. The confusion matrix in Table 2 shows that anger and joy are identified well with an accuracy of over 90%, while other five expressions are discriminated poorly with an accuracy of less than 90%. In particular, sad is classified with the lowest accuracy of 64.93% since sad is highly confused to fear and neutral.

Experiments on the Cohn-Kanade database

The Cohn-Kanade database [29] consists of 100 university students aged from 18 to 30 years. Image sequences from neutral to target display were digitized into 640*490 pixels with 8-bit precision for grayscale values. As done in [11,12], 320 image sequences were selected from 96 subjects for experiments. For each sequence, the neutral face and three peak frames were used for prototypic expression recognition, resulting in 1409 images (anger 96, joy 298, sad 165, surprise 225, fear 141, disgust 135, and neutral 349). Figure 4 shows a few

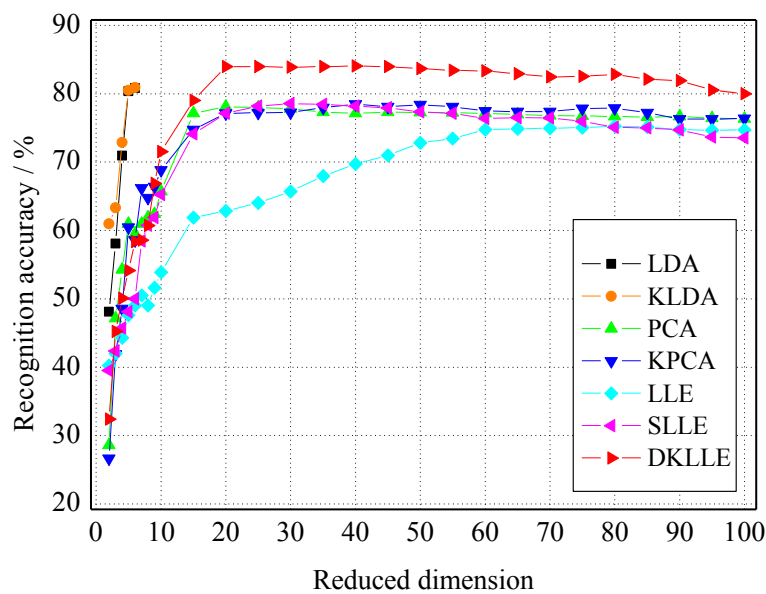


Figure 3 Recognition accuracy vs. reduced dimension on the JAFFE database.

Table 1 The best accuracy (std) of different methods on the JAFFE database

Method	LDA	PCA	KLDA	KPCA	LLE	SLLE	DKLLE
Dimension	6	20	6	40	80	30	40
Accuracy (%)	80.81 ± 3.6	78.09 ± 4.2	80.93 ± 3.9	78.47 ± 4.0	75.24 ± 3.8	78.57 ± 4.0	84.06 ± 3.8

Table 2 Confusion matrix of recognition results with DKLLE on the JAFFE database

	Anger (%)	Joy (%)	Sad (%)	Surprise (%)	Disgust (%)	Fear (%)	Neutral (%)
Anger	91.85	0	3.11	0	2.80	0	2.24
Joy	0	95.20	2.27	0	0	0	2.53
Sad	5.82	2.96	64.93	0.02	3.00	9.00	14.27
Surprise	0	3.05	2.39	89.16	0	5.32	0.08
Disgust	6.75	0	2.63	0	83.62	7.00	0
Fear	0	0.03	11.14	5.98	2.10	80.75	0
Neutral	0	0	15.87	1.22	0	0	82.91

Values in boldface represent accuracy per expression.



Figure 4 Examples of facial expression images from the Cohn-Kanade database.

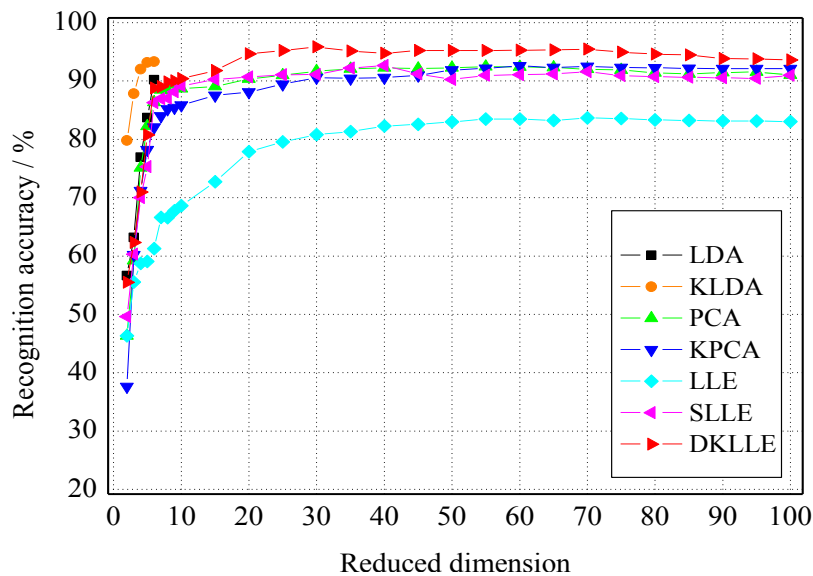


Figure 5 Recognition accuracy vs. reduced dimension on the Cohn-Kanade database.

Table 3 The best accuracy (std) of different methods on the Cohn-Kanade database

Method	LDA	PCA	KLDA	KPCA	LLE	SLLE	DKLLE
Dimension	6	55	6	60	70	40	30
Accuracy (%)	90.18 ± 3.0	92.43 ± 3.3	93.32 ± 3.0	92.59 ± 3.6	83.67 ± 3.4	92.64 ± 3.2	95.85 ± 3.2

Table 4 Confusion matrix of recognition results with DKLLE on the Cohn-Kanade database

	Anger (%)	Joy (%)	Sad (%)	Surprise (%)	Disgust (%)	Fear (%)	Neutral (%)
Anger	98.23	0	0.77	0	0	1.00	0
Joy	0.25	96.78	0.20	0	1.38	0.25	1.14
Sad	1.78	0.94	91.12	0	5.16	0	1.00
Surprise	0.20	0.20	1.43	97.84	0	0	0.33
Disgust	0	0.73	1.00	2.76	95.19	0.32	0
Fear	0	0	0	0.15	0	99.85	0
Neutral	2.00	1.13	2.96	0.42	1.13	0.42	91.94

Values in boldface represent accuracy per expression.

examples of facial expression images from the Cohn-Kanade database.

Figure 5 presents the recognition performance of different methods. Table 3 shows the best accuracy (std) for different methods with the corresponding reduced dimension. The results in Figure 5 and Table 3 indicate that DKLLE obtains the recognition performance superior to the other used methods again. Compared with the previously reported work [11,12] in which the experimental settings are similar to ours, the best accuracy of 95.85% obtained by LBP-based 1-NN is highly competitive. In [11], on 7-class facial expression recognition tasks they used LBP-based template matching and reported an accuracy of 79.1%. Additionally, they also employed LBP-based SVM to give an accuracy of 87.2, 88.4, and 87.6% with linear, polynomial and RBF kernels, respectively. In [12], based on boosted-LBP features and SVM, on 7-class facial expression recognition tasks they reported an accuracy of 91.1, 91.1, and 91.4% with linear, polynomial and RBF kernels, respectively.

Table 4 shows the confusion matrix of 7-class expression recognition results when DKLLE obtains the best performance at 30 reduced dimension. From Table 4, it can be seen that 7-class facial expressions are identified very well with an accuracy of over 90%.

Conclusions

A new kernel-based supervised manifold learning algorithm, called DKLLE, is proposed for facial expression recognition. DKLLE has two prominent characteristics. First, as a kernel-based feature extraction method, DKLLE can extract the nonlinear feature information embedded on a data set, as KPCA and KLDA does. Second, DKLLE is designed to obtain a high discriminating power for its low-dimensional embedded data representations in an effort to improve the performance on facial

expression recognition. Experimental results on the JAFFE database and the Cohn-Kanade Database show that DKLLE not only makes an obvious improvement over LLE and SLLE, but also outperforms the other used methods including PCA, LDA, KPCA, and KLDA.

Acknowledgements

This work was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. Z1101048 and Grant No. Y1111058.

Author details

¹Department of Computer Science, Taizhou University, Taizhou 318000, P.R. China ²School of Physics and Electronic Engineering, Taizhou University, Taizhou 318000, P.R. China

Competing interests

The authors declare that they have no competing interests.

Received: 4 October 2011 Accepted: 27 January 2012

Published: 27 January 2012

References

1. RW Picard, *Affective Computing*, (The MIT Press, Cambridge, 2000)
2. Y Tian, T Kanade, J Cohn, *Facial expression analysis, Handbook of face recognition*, (Springer, Heidelberg, 2005), pp. 247–275
3. P Viola, M Jones, Robust real-time face detection. *Int J Comput Vision*. **57**(2), 137–154 (2004)
4. MA Turk, AP Pentland, Face recognition using eigenfaces, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society, HI USA, pp. 586–591 (1991)
5. PN Belhumeur, JP Hespanha, DJ Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell*. **19**(7), 711–720 (1997). doi:10.1109/34.598228
6. CC Lee, SS Huang, CY Shih, Facial affect recognition using regularized discriminant analysis-based algorithms. *EURASIP J Adv Signal Process*. **2010**, 10 (2010)
7. JG Daugman, Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Trans Acoust Speech Signal Process*. **36**(7), 1169–1179 (1988). doi:10.1109/29.1644
8. L Shen, L Bai, Information theory for Gabor feature selection for face recognition. *EURASIP J Adv Signal Process*. **2006**, 11 (2006)
9. T Ahonen, A Hadid, M Pietikäinen, Face description with local binary patterns: Application to face recognition. *IEEE Trans Pattern Anal Mach Intell*. **28**(12), 2037–2041 (2006)

10. T Ojala, M Pietikäinen, T Mäenpää, Multiresolution gray scale and rotation invariant texture analysis with local binary patterns. *IEEE Trans Pattern Anal Mach Intell.* **24**(7), 971–987 (2002). doi:10.1109/TPAMI.2002.1017623
11. C Shan, S Gong, P McOwan, Robust facial expression recognition using local binary patterns, in *IEEE International Conference on Image Processing (ICIP)*, vol. 2. IEEE Computer Society, Italy, pp. 370–373 (2005)
12. C Shan, S Gong, P McOwan, Facial expression recognition based on local binary patterns: a comprehensive study. *Image Vis Comput.* **27**(6), 803–816 (2009). doi:10.1016/j.imavis.2008.08.005
13. S Moore, R Bowden, Local binary patterns for multi-view facial expression recognition. *Comput Vis Image Understand.* **115**, 541–558 (2011). doi:10.1016/j.cviu.2010.12.001
14. Y Tian, T Kanade, J Cohn, Recognizing action units for facial expression analysis. *IEEE Trans Pattern Anal Mach Intell.* **23**(2), 97–115 (2002)
15. MJ Lyons, J Budynek, S Akamatsu, Automatic classification of single facial images. *IEEE Trans Pattern Anal Mach Intell.* **21**(12), 1357–1362 (1999). doi:10.1109/34.817413
16. N Sebe, MS Lew, Y Sun, I Cohen, T Gevers, TS Huang, Authentic facial expression analysis. *Image Vis Comput.* **25**(12), 1856–1863 (2007). doi:10.1016/j.imavis.2005.12.021
17. I Kotsia, I Pitas, Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE Trans on Image Process.* **16**(1), 172–187 (2007)
18. Y Chang, C Hu, M Turk, Manifold of facial expression, in *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, IEEE Computer Society, France, pp. 28–35 (2003)
19. C Shan, S Gong, PW McOwan, Appearance manifold of facial expression, in *Computer Vision in Human-Computer Interaction, Lecture Notes in Computer Science*, vol. 3766. Springer, China, pp. 221–230 (2005). doi:10.1007/11573425_22
20. Y Chang, C Hu, R Feris, et al, Manifold based analysis of facial expression. *Image Vis Comput.* **24**(6), 605–614 (2006). doi:10.1016/j.imavis.2005.08.006
21. ST Roweis, LK Saul, Nonlinear dimensionality reduction by locally linear embedding. *Science.* **290**(5500), 2323–2326 (2000). doi:10.1126/science.290.5500.2323
22. JB Tenenbaum, V de Silva, JC Langford, A global geometric framework for nonlinear dimensionality reduction. *Science.* **290**(5500), 2319–2323 (2000). doi:10.1126/science.290.5500.2319
23. Y Cheon, D Kim, Natural facial expression recognition using differential-AAM and manifold learning. *Pattern Recogn.* **42**(7), 1340–1350 (2009). doi:10.1016/j.patcog.2008.10.010
24. R Xiao, Q Zhao, D Zhang, P Shi, Facial expression recognition on multiple manifolds. *Pattern Recogn.* **44**(1), 107–116 (2011). doi:10.1016/j.patcog.2010.07.017
25. D de Ridder, O Kouropteva, O Okun, M Pietikäinen, RPW Duin, Supervised locally linear embedding, in *Artificial Neural Networks and Neural Information Processing-ICANN/ICONIP-2003, Lecture Notes in Computer Science*, vol. 2714. Springer, Heidelberg, pp. 333–341 (2003). doi:10.1007/3-540-44989-2_40
26. L Zhao, Z Zhang, Supervised locally linear embedding with probability-based distance for classification. *Comput Math Appl.* **57**(6), 919–926 (2009). doi:10.1016/j.camwa.2008.10.055
27. B Li, C-H Zheng, D-S Huang, Locally linear discriminant embedding: An efficient method for face recognition. *Pattern Recogn.* **42**(12), 3813–3821 (2008)
28. D Liang, J Yang, Z Zheng, Yuchou Chang, A facial expression recognition system based on supervised locally linear embedding. *Pattern Recogn Lett.* **26**(15), 2374–2389 (2005). doi:10.1016/j.patrec.2005.04.011
29. T Kanade, Y Tian, J Cohn, Comprehensive database for facial expression analysis, in *International Conference on Face and Gesture Recognition*, vol. 4. IEEE Computer Society, France, pp. 46–53 (2000)
30. B Scholkopf, The kernel trick for distances, in *Advances in Neural Information Processing Systems*, (MIT Press Cambridge, Canada, 2001), pp. 301–307
31. B Scholkopf, A Smola, K Muller, Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **10**(5), 1299–1319 (1998). doi:10.1162/089976698300017467
32. G Baudat, F Anouar, Generalized discriminant analysis using a kernel approach. *Neural comput.* **12**(10), 2385–2404 (2000). doi:10.1162/089976600300014980
33. J Wang, Z Zhang, H Zha, Adaptive manifold learning, in *Advances in neural information processing systems*, vol. 17. MIT Press Cambridge, Canada, 2005, pp. 1473–1480
34. LK Saul, ST Roweis, Think globally, fit locally: unsupervised learning of low dimensional manifolds. *J Mach Learn Res.* **4**, 119–155 (2003)
35. P Campadelli, R Lanzarotti, G Lipori, E Salvi, Face and facial feature localization, in *International Conference on Image Analysis and Processing*, Springer, Heidelberg, Italy, pp. 1002–1009 (2005)

doi:10.1186/1687-6180-2012-20

Cite this article as: Zhao and Zhang: Facial expression recognition using local binary patterns and discriminant kernel locally linear embedding. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:20.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
