# Face recognition using data driven local appearance features

Xin Xie, Chao Chen and Zhijian Chen

School of Zhejiang University, Hangzhou, 310027, China

xiexin@zju.edu.cn, chenzj@vlsi.zju.edu.cn

**Keywords:** Face recognition, local appearance features, data driven, feature extraction, point-to-subspace distance.

**Abstract.** A novel data driven face descriptor based on point-to-subspace metric is proposed for subspace classifiers. Unlike conventional feature descriptors which are carefully designed by hand, the newly proposed method uses supervised learning to derive more robust and more discriminative descriptors. During the feature extraction process, the point-to-subspace distance is used as the inner mechanism to train parameters including filters and weights of different pixels. Experimental results on FERET and Extended Yale B database show that when using subspace classifiers, the proposed feature descriptor is more discriminative and yields higher recognition rate over other features.

## Introduction

Face recognition involves two main parts: feature representation and classifier selection. For the last decades, both of the two aspects have been the hotspots in the research field, in order to improve the facial recognition performance.

Among various classification methods, the nearest neighbor (NN) [1], Adaboost [2] and support vector machine (SVM) [3] are the most popular classifiers in face recognition. These methods compare the point-to-point distances between the gallery and probe sets. Recently, subspace classifiers, such as sparse representation classifier (SRC) [4] and linear regression classifier (LRC) [5], have attracted significant attention due to their outstanding performance. Theoretically these methods measure the point-to-subspace distance [6] following the basic principle that images from the same person lie in the same subspace. Fig. 1 gives the geometric interpretation of these two kinds of distances.

Feature representation is another important aspect of face recognition. It is defined as the procedure that attempts to extract the most discriminative elements to represent the original images. Various features have been used to represent faces, e.g. holistic features, such as principal component analysis (PCA) [7] and linear discriminant analysis (LDA) [8], or local appearance features, such as local binary patterns (LBP) [9] and local quantized patterns (LQP) [10]. All of these features are designed in a hand-crafted way. Recently, to generate more robust and discriminative features, data driven methods have been adopted. Most of the existing works attempt to combine data driven features with point-to-point classifiers. For instance, Ren [11] uses the locality principle based on two elements to derive a set of highly discriminative features for facial landmark representation.

However, when using subspace classifiers, most of the previous works [4, 5] simply combine existing feature descriptors with SRC or LRC, while ignoring the inner classification mechanism of subspace classifiers. In this paper, a novel data driven feature descriptor is specially designed for subspace classifiers. This newly proposed feature follows the point-to-subspace metric when learning filters and weights of image patches. Experimental results exhibit that the descriptor has more discriminative ability and achieves higher recognition rates.

 

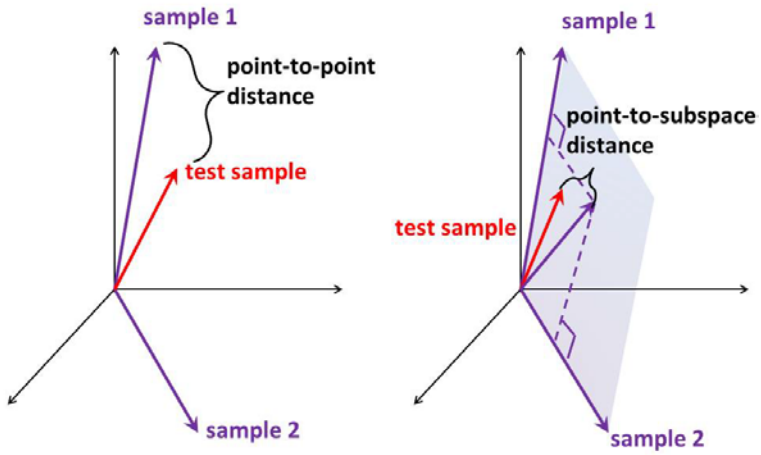


Fig. 1 An illustration of the point-to-point distance and the point-to-subspace distance

## Data driven local appearance features

Local appearance features have certain advantages when meeting small changes such as occlusion and expression. The proposed descriptor contains the basic components of local appearance features and adds optimization process to select the most proper parameters. A schematic of the procedure is shown in Fig. 2.

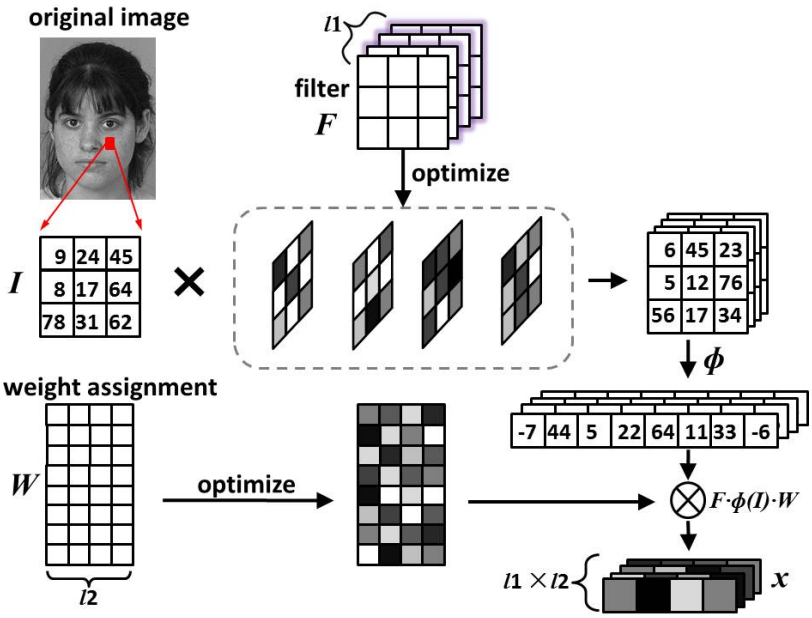


Fig. 2 An example of the feature extraction procedure. The filter contains $l1$ complemental parts of size 3*3, the weight matrix size is 8*$l2$, and the final descriptor has a size of $l1$*$l2$.

Given a face image $I$, a filter is first applied to reduce the noise and enhance the useful information. The filtered image is denoted $f(I)$. Then, the LBP-like operation $\varphi$ is adopted which compares the neighboring pixels with the central one. For position $p_0$ in the original image, the operation on the filtered image can be expressed as $\varphi(f(p_0))=[f(p_1)-f(p_0),f(p_2)-f(p_0),\ldots,f(p_c)-f(p_0)]$, where $p_1,p_2,\ldots,p_c$ are the pixels around $p_0$ and $c$ is the number of selected neighbor points. Because different neighbors may have different contributions to the final recognition performance, a weight assignment step is added.

Using a linear assumption, the filters can be represented by a matrix $F$; thus, $f(I)=F \bullet I$. The weight assignment step can also be completed by multiplying a weight matrix $W$. Then, the entire generation of the proposed feature matrix can be expressed as $\varphi(F \bullet I) \bullet W=F \bullet \varphi(I) \bullet W$.

To precisely determine the values of the filters and the weight assignment matrix, data driven methods are applied and the gallery image sets are used for learning parameters. Fig.2 points out the two steps that need to be optimized. During optimization, a point-to-subspace distance is used, and this metric is specifically designed for subspace classifiers such as SRC and LRC. An illustration of the point-to-subspace distances is already given in Fig. 1. The basic idea behind optimization is that after filtering and weight assignment, the images of the same person are more likely to lie in the same subspace, and the distances between the subspaces for different people are increased.

For a processed image $x_{i,j}=F \bullet \varphi(I_{i,j}) \bullet W$, where $i,j$ denotes the $j_{th}$ sample from class $i$. The point-to-subspace distance in the same person is calculated as

$$\begin{aligned} D_{intra}(x_{i,j}) &= x_{i,j} - X_{i,\bar{j}} \cdot \alpha \\ &= \mathrm{F} \cdot (\varphi(I_{i,j}) - \varphi(I_{i,\bar{j}}) \cdot \alpha) \cdot W \\ &= \mathrm{F} \cdot d_{intra}(x_{i,j}) \cdot W \end{aligned} \tag{1}$$

where $X_{i,\bar{j}}$ is a matrix composed of all of the training samples except the $j_{th}$ of class $i$. $\alpha$ is the coefficient, which could be obtained using linear regression. Because both $X_{i,j}$ and $X_{i,\bar{j}}$ contain $F$ and $W$, they can be extracted and expressed in a simple way as the last line in (1). Similarly, the inter-class distance of $x_{i,j}$ with respect to class $r$ $(r \neq i)$ is defined as

$$\begin{aligned} D_{inter}^{r}(x_{i,j}) &= x_{i,j} - X_r \cdot \beta \\ &= \mathrm{F} \cdot d_{inter}^{r}(x_{i,j}) \cdot W \end{aligned} \tag{2}$$

where $X_r$ is composed of all of the training samples from the $r_{th}$ class.

Through optimization procedure, the proper $F$ and $W$ are calculated aiming to achieve the maximum ratio of the inter-class distance and the intra-class distance. The procedure is formulated as

$$\begin{aligned} max\ J(F,W) &= \frac{\sum_r \sum_i \sum_j D_{inter}^{r}(x_{i,j})}{\sum_i \sum_j D_{intra}(x_{i,j})} \\ &= \frac{F \cdot (\sum_r \sum_i \sum_j d_{inter}^{r}(x_{i,j})) \cdot W}{F \cdot (\sum_i \sum_j d_{intra}(x_{i,j})) \cdot W} \end{aligned} \tag{3}$$

The problem in (3) can be solved in an iterative way. One of the two parameter matrices is fixed, and the other one is derived by solving the generalized eigenvalue problem.

After the $F$ and $W$ matrices are obtained, the descriptors with size $l1*l2$ are re-calculated. Then, an unsupervised learning method, such as $K$-means, is applied to extract the dominant patterns, which could enhance the effectiveness of the descriptors. Additionally, to further reduce the dimension of the vector, the holistic projection step, such as PCA or LDA, can be added as the last step of the feature generation process if necessary. Also, the metric used for holistic projection should be the point-to-subspace distance.

## Experimental results

The experiments are performed on the FERET and Extended Yale B datasets. The FERET subset used in this letter contains 1400 images from 200 individuals with 7 images of each person. All of the images are cropped to a size of 40*40. The Extended Yale B dataset is composed of 2414 frontal face images from 38 subjects under different lighting conditions. All of the images are downsampled to a size of 30*30. All our experiments are tested on Matlab R2012a.

In our experiment, the recognition rates of various feature descriptors, including PCA [7], LDA [8], LBP [9], CCEDA [6] and our proposed method are evaluated and compared on the two databases, respectively. The number of dimensions for all features is restricted to 400. Table 1 shows the recognition rates under different subsets on FERET. The results show that the proposed method can improve the performance significantly. The Extended Yale B database was divided into 5 subsets. Subset 1 is used as the gallery and contains 266 images under nominal lighting conditions.

The remaining subsets are used as the probe sets. Among them, subsets 2 and 3 have slight-to-moderate luminance variations, while subsets 4 and 5 have severe light variations. The results are listed in Table 2, and also demonstrate that the proposed data driven descriptors outperform the other descriptors.

Table 1 The recognition performance on FERET database using SRC

| method | Fb (%) | Fc (%) | Dup I (%) | Dup II (%) |
|---|---|---|---|---|
| PCA [7] | 74.2 | 70.8 | 65.3 | 60.3 |
| LDA [8] | 76.9 | 76.1 | 68.2 | 62.7 |
| LBP [9] | 86.5 | 83.9 | 75.4 | 73.2 |
| CCEDA [6] | 92.6 | 89.4 | 77.2 | 76.1 |
| Proposed | **99.1** | **98.2** | **83.7** | **80.5** |

Table 2 The recognition performance on Extended Yale B database using LRC

| method | Subset 2 (%) | Subset 3 (%) | Subset 4 (%) | Subset 5 (%) |
|---|---|---|---|---|
| PCA [7] | 97.50 | 78.30 | 15.10 | 23.25 |
| LDA [8] | 97.65 | 80.40 | 14.25 | 23.90 |
| LBP [9] | 98.10 | 77.50 | 15.67 | 26.25 |
| CCEDA [6] | 98.30 | 90.20 | 30.50 | 27.10 |
| Proposed | **100** | **100** | **80.10** | **31.20** |

## Conclusion

In this paper, a new data driven feature descriptors specifically designed for subspace classifiers including SRC and LRC is proposed. The point-to-subspace distance is used as the criterion to generate more discriminative descriptors. After adding the optimized parameters, the discriminative ability of the proposed descriptors is greatly improved. The experimental results on two popular databases confirm that our proposed features achieve satisfactory performance.

## Reference


[1]. Chien J T, Wu C C. Discriminant waveletfaces and nearest feature classifiers for face recognition. Pattern Analysis & Machine Intelligence IEEE Transactions on, Vol. 24 (2002) No. 12, p. 1644-1649.

[2]. Vapnik V. Statistical Learning Theory. Wiley-Interscience, 1998.

[3]. Bishop C M. Pattern Recognition and Machine Learning. Journal of Electronic Imaging, Vol. 16 (2006) No. 4.

[4]. Wright J, Member S, Yang A Y, et al. Robust Face Recognition Via Sparse Representation. IEEE Trans. Pattern Analysis and Machine Intelligence. Vol. 31 (2009), p. 210-227.

[5]. Naseem I, Togneri R, Bennamoun M. Linear Regression for Face Recognition. Pattern Analysis & Machine Intelligence IEEE Transactions on. Vol. 32 (2010) No. 11, p. 2106-2112.

[6]. Zhu Q, Sun H, Feng Q, et al. CCEDA: building bridge between subspace projection learning and sparse representation-based classification. Electronics Letters. Vol. 50 (2014) No. 25, p. 1919-1921.

[7]. M.A. Turk, A.P. Pentland. Face Recognition Using Eigenfaces. Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 1991, p. 586-591.

[8]. P.N. Belhumeur, J.P. Hespanha and D.J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Trans. Pattern Analysis and Machine Intelligence. Vol. 19 (1997) No. 7, p. 711-720.

[9]. T. Ahonen, A. Hadid and M. Pietikainen. Face Description with Local Binary Patterns: Application to Face Recognition. IEEE Trans. Pattern Analysis and Machine Intelligence. Vol. 28 (2006) No. 12, p. 2037-2041.

[10]. S. ul Hussain, T. Napoleon and F. Jurie. Face Recognition Using Local Quantized Patterns. Proc. British Machine Vision Conf. (BMVC), 2012.

[11]. Shaoqing Ren, Xudong Cao, Yichen Wei, et al.: Face alignment at 3000 FPS via regressing local binary features. IEEE Conf. Computer Vision and Pattern Recognition, 2014, pp. 1685-1692