Multimedia Tools and Applications (2020) 79:675–691
https://doi.org/10.1007/s11042-019-08157-0

Face recognition based on local binary pattern and improved Pairwise-constrained Multiple Metric Learning

Lijian Zhou[1] · Hui Wang[1,2] · Shanshan Lin[1] · Siyuan Hao[1] · Zhe-Ming Lu[3]

Received: 17 September 2018 / Revised: 3 July 2019 / Accepted: 2 September 2019
Published online: 7 September 2019

Abstract
During the image acquisition process of face recognition, the obtained face images are affected inevitably by varied illumination and position in different environment. Local Binary Pattern (LBP) operator is used to decrease illumination effectiveness. Improved Pairwise-constrained Multiple Metric Learning method (IPMML) is proposed as a classification metric in our prior work, which solves the misalignment problem in a better way compared with PMML. To solve the high computation complexity of IPMML, Linear Discriminant Analysis (LDA) is performed before IPMML. Thus, a face recognition method based on LBP and IPMML is proposed, which can overcome the illumination and misalignment problems. LBP is selected to extract texture features of face images firstly. Second, LDA is applied to reduce the dimension. Then the fisher features are divided into sub-blocks according to the dimension of features and every block is a column vector. Fourth, a classification metric – IPMML is used to obtain the optimum Mahalanobis matrix. Fifth, the Mahalanobis matrix is used to compute the final discriminative distance. Finally, the Nearest Neighborhood Classifier (NNC) is applied to classify face images. The experimental results show that the proposed method can achieve high recognition rates and is robust to illumination and facial expression variation, especially for misaligned face images.

Keywords Face Recognition · LBP · LDA · IPMML · NNC

✉ Lijian Zhou
zhoulijian@qut.edu.cn

Hui Wang
972482380@qq.com

Zhe-Ming Lu
zheminglu@zju.edu.cn

Extended author information available on the last page of the article

1 Introduction

In face recognition system, the input face images often contain variable illumination and noise due to the diverse acquisition environment. LBP [8] is a gray-scale invariant texture description operator which can weaken the influence of variable illumination. By extracting the LBP feature, the distinguishing ability is improved since the texture information is enhanced. Li et al. [15] obtained the first-order differential feature using LBP and the second order differential feature using Center-Symmetric Local Derivative Pattern (CS-LDP), and then fused them to get richer texture information. Zhu et al. [21] got the LBP/LPQ histograms by fusing LBP feature in the spatial domain and Local Phase Quantization (LPQ) feature in the frequency domain. Wang et al. [13] extracted the statistical characteristic histograms of image blocks using different operators. The histograms are then linked to get the CLBP features which are used to describe the face. Although the three approaches mentioned above can extract rich texture features, the dimension of the extracted features is larger than that of original face images. LDA [1] is an effective dimensional reduction method and the obtained feature contains class information to improve the recognition accuracy. Therefore, LDA is usually used to decrease feature dimension and reduce computing time. Dictionary learning and deep learning methods [4, 11, 12, 14] have been used more frequently. Wang *et al*. [9] used LeNET Convolutional Neural Networks (CNN) to recognize face. He et al. [10] proposed a partial face recognition method using fully convolutional network and sparse representation classification. Wang *et al*. [5] proposed a large margin cosine loss for deep face recognition, which can achieve the minimum intra-class variance and maximum inter-class variance by virtue of normalization and cosine decision margin maximization. References [6, 17, 18] proposed the face recognition based on LBP and CNN. They can effectively recognize the face, but a large number of the training samples are needed, which consumes much time and large computation resources. Nearest Neighborhood Classifier (NNC) [19] is a universally classification method since it is easy to realize. The Euclidean distance is the most commonly used discriminative metric in NNC. However, it needs accurate alignment between testing and training samples. Cui et al. [2] proposed a new classification metric – PMML which computes the Mahalanobis matrix between the testing sample blocks and the whole training sample blocks. Nevertheless, it does not do well in recognizing misaligned face images. IPMML [20] is proposed to compute the optimum Mahalanobis matrices between the testing face image and every training face image. By comparing the above two classification metrics, IPMML can preserve the detailed information of face images as well as recognize misaligned face images effectively. To improve the recognition result and decrease the computation, the fisher features are extracted using the LDA method in our prior paper [20]. By observing the recognition results, it is observed that the misidentification occurs with high probability while the illumination, pose and appearance vary hugely from the training samples. Comparing with low and middle frequency features of wavelet, multiwavelet and curvelet, LBP features preserve more detailed information. Therefore, a new face recognition approach based on LBP and IPMML is proposed. First, the features of face images are extracted based on LBP and LDA. Next, the final features with low dimension are divided into a number of blocks. And then, the final discriminative distance is computed using the IPMML method. Finally, NNC is performed to classify the face images. Experiments on Yale, Extended Yale B and AR Databases show that the proposed method can achieve high recognition rate and robustness.

The rest of this paper is organized as follows. In Section 2, we give the proposed method. Then, we do experiments on Yale, Extended Yale B and AR Databases and analyze the results in Section 3. Finally, some conclusions are obtained in Section 4.

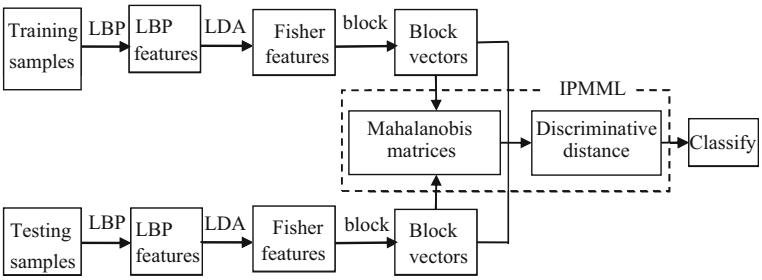


Fig. 1 The face recognition framework based on LBP and IPMML

2 Proposed face recognition method

To classify face images better, a face recognition approach based on LBP and IPMML is proposed in this paper. LBP is a texture descriptor which can extract texture features effectively and has strong distinguishing ability. LDA is a commonly used dimension reduction method. IPMML is applied to obtain the optimized Mahalanobis matrix and achieve a better recognition rate. The framework of the proposed method is shown in Fig. 1. First, LBP operator is applied to extract texture features of original images as the elementary features. Second, the fisher features are obtained using the LDA method as the final features. Third, the final features are divided into the blocks according to the feature dimension and every column is regarded as an image block. Fourth, the Mahalanobis matrix is obtained between the testing sample blocks and the training sample blocks. Fifth, compute the discriminative distance using the Mahalanobis matrix. Finally, classify face images according to the theory of NNC. The detailed process of the proposed approach is as follows.

2.1 The elementary face feature extraction based on LBP

LBP [8] operator is a gray-scale invariant texture description operator which is insensitive to light changes. The calculation of LBP is very simple and it does not need to set parameters in advance. Thus, LBP is one of the most widely used texture description operators.

The basic LBP operator is used in a 3×3 window as shown in Fig. 1a. The central pixel g_c is considered as the threshold, and the other eight pixels g_i, $i(0 \le i \le 7)$ are the adjacent pixels. The

(a) Original images

(b) LBP images of (a)

Fig. 2 Original images and its LBP images

Table 1 The fisher features of the LBP features and blocking

	The final features													
S11	−492.61	585.75	745.75	−1024.81	−1603.49	−9.38	990.41	179.83	−651.68	−732.16	290.54	360.30	355.83	34.20
S12	−494.02	587.35	770.03	−1040.36	−1592.53	5.27	1047.69	251.07	−648.79	−644.96	388.06	198.75	312.52	341.21
S13	−490.71	578.98	778.73	−1054.67	−1561.96	31.26	983.28	429.15	−707.11	−798.71	346.91	126.09	420.04	358.61
S14	−500.97	579.38	731.14	−1056.32	−1515.96	−39.10	1107.73	166.42	−807.17	−562.80	164.19	318.78	−15.66	37.68
S15	−493.98	571.92	774.58	−1109.88	−1587.26	−18.05	1125.35	365.60	−854.92	−618.46	163.41	194.79	142.86	174.59
S21	−91.79	1513.16	−568.15	−52.49	−2032.52	−1878.91	23.16	1237.85	749.13	736.07	−424.29	−281.45	0.16	−145.94
S22	−87.76	1503.16	−547.59	−61.13	−1973.66	−2008.32	77.32	1105.06	600.57	793.54	−831.36	−354.00	300.21	−272.74
S23	−92.86	1506.63	−555.41	−53.74	−2023.37	−1953.18	−1.35	1058.34	778.19	778.55	−645.23	−202.03	423.19	−146.24
S24	−90.09	1507.13	−582.97	−67.18	−1979.18	−1971.56	114.17	1201.93	564.04	753.06	−857.04	−457.69	204.04	−536.07
S25	−91.96	1514.23	−590.43	−64.04	−1943.91	−1989.25	55.35	966.53	750.47	830.48	−749.84	−193.71	104.24	−141.96

Table 2 The fisher features of the LBP features and blocking of the testing sample

	The final features													
T11	−435.95	94.22	1029.14	−1174.57	−1813.88	−500.53	457.25	176.24	94.00	−38.09	−385.32	−144.89	119.97	−74.06

Table 3 The distance between the first testing sample and every training sample used the IPMML method ($\times 10^6$)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
1	0.1368	2.61	2.97	4.09	3.47	3.55	4.19	3.59	8.02	3.56	6.94	3.28	3.52	5.34	5.13
2	0.1369	2.60	2.92	4.09	3.41	3.41	4.10	3.60	7.98	3.65	6.97	3.25	3.56	5.16	5.18
3	0.1364	2.54	2.95	4.09	3.48	3.85	3.95	3.56	7.99	3.41	6.80	3.25	3.51	5.62	5.03
4	0.1367	2.56	2.95	4.12	3.42	3.59	4.17	3.59	7.93	3.37	6.92	3.33	3.52	5.33	5.09
5	0.1367	2.70	2.91	4.09	3.44	3.51	4.11	3.56	8.04	3.83	6.88	3.29	3.55	5.33	5.01

adjacent pixel value is denoted as 1 if it is larger than the threshold value, otherwise 0. Then the texture value T is computed according to eq. (1).

The texture value of point (x_c,y_c) is computed as follows.

$$T(x_c,y_c)=\sum_{i=0}^{7} s(g_i-g_c)2^i \tag{1}$$

where, $s(x)=\begin{cases}1 & x>0\\ 0 & x\le 0\end{cases}$, g_c is the central pixel value, g_i, $i(0\le i\le 7)$ is the adjacent pixel value.

Five face images (with the size of 64×64) of a person are chosen from Yale database (Fig. 2a) as an example and perform the LBP operation on it. The result is shown in Fig. 2b, whose size is 62×62. It is obvious that the LBP image not only can reflect texture feature of all critical regions clearly, but also weaken the feature of some smooth regions comparing with the original images. In addition, the illumination effects can be removed significantly at the same time.

2.2 The final fisher features extraction and blocking

Every LBP feature image of the training samples obtained using the LBP operation is reshaped to a vector. Then the fisher feature is computed using the LDA method, which is a vector. At last, the vector is divided to the sub-blocks with an equal size. Five images of everyone are taken as the training samples. Every LBP feature image is reshaped to a vector with the size of 1×3844. Take the images in Fig. 2 as an example. Their fisher features are shown in Table 1 (the frontier 5 rows), where S*ij* represents the *j* training samples of the *i*th person. To observe clearly, the features of the second person are given in the 6th -10th row at the same time.

In general, the dimension of the final fisher feature is small and the size of every block should be integer, so the size of the block is determined to the least factor of the dimension of the final feature except 1. Since the dimension of the final feature is 14 in this example, every vector is divided to 7 blocks with size of 1×2. It can be seen that the features of the samples with same identification are similar. The differences are obvious for different identifications, so the extracted features using the LBP and LDA [16] methods are effective.

2.3 The classification based on IPMML

PMML [2] is used to compute a Mahalanobis matrix which is applied to get the discriminative distance. The Mahalanobis matrix is obtained according to the testing sample and the whole training samples. For better recognizing misalignment samples, IPMML [20] is proposed in our prior work to obtain the Mahalanobis matrix between the testing sample and the training sample, and then the final distance is computed using the Mahalanobis matrix. As a classification method, the most important thing of IPMML is to obtain an optimized Mahalanobis matrix.

Table 4 The distance between the first testing sample and every training sample used the PMML method ($\times 10^6$)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
1	0.99	10.28	9.94	10.88	13.21	12.03	13.04	13.30	12.83	16.61	14.71	11.93	10.92	11.27	10.92
2	1.06	10.19	11.61	10.11	11.49	10.70	10.21	13.49	10.22	15.98	12.40	11.42	11.82	10.37	9.57
3	1.06	10.75	10.52	11.83	12.15	11.31	10.86	13.40	10.10	15.11	12.13	11.29	12.51	10.10	10.22
4	1.17	10.22	10.23	10.71	12.24	10.67	11.03	13.28	10.54	15.75	12.52	11.95	12.14	9.91	9.95
5	1.36	10.04	10.31	10.91	12.12	10.48	10.04	13.78	10.47	15.45	13.55	11.79	11.87	10.74	9.77

Table 5 The Euclidean distance between the first testing sample and every training sample ($\times 10^3$)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
1	1.40	2.75	2.71	3.40	3.07	3.02	3.19	3.21	3.37	3.19	3.23	2.95	3.05	3.33	3.02
2	1.47	3.11	2.65	3.40	3.09	3.19	2.98	3.24	3.26	3.37	3.19	2.98	3.23	3.16	3.02
3	1.57	2.79	2.87	3.42	3.14	3.38	3.32	3.09	3.53	3.56	3.05	3.02	2.82	3.60	3.01
4	1.52	2.78	2.73	3.34	3.10	3.21	3.28	3.16	3.65	3.17	3.13	3.07	3.14	3.33	3.02
5	1.44	3.06	2.62	3.25	3.17	3.08	3.17	3.19	3.19	3.35	3.14	2.98	3.17	3.32	2.99

Suppose training face image $\mathbf{B}_i^k (i = 1, \cdots, N, k = 1, , \cdots K)$ and testing face image $\mathbf{B}_j^k$ $(j = 1, \cdots, M, k = 1, , \cdots K)$ are the k-th block of the i-th sample and the j-th sample separately, then the distance between them is computed as follows.

$$d(\mathbf{B}_j, \mathbf{B}_i) = \frac{1}{2} \sum_{k=1}^{K} \left(\mathbf{B}_i^k - \mathbf{B}_j^k\right)^T \mathbf{W}_k \left(\mathbf{B}_i^k, \mathbf{B}_j^k\right) \tag{2}$$

where, $\mathbf{W}_k$ is the Mahalanobis matrix of the k-th block. For two given samples $\mathbf{B}_j$ and $\mathbf{B}_i$, $d(\mathbf{B}_j,\mathbf{B}_i)$ need to be smaller than the certain threshold ρ when the two samples belong to the same person. Otherwise, $d(\mathbf{B}_j,\mathbf{B}_i)$ needs to be larger than ρ.

To obtain the optimal Mahalanobis matrix $\mathbf{W}_k$, the objective function of IPMML is defined as follows.

$$\min_{\mathbf{W}_k, \xi_{ij}} \frac{1}{K} \sum_{k=1}^{K} H(\mathbf{W}_k, \mathbf{W}_0) + \frac{\gamma}{n} \sum_{i,j} l\left(\xi_{ij}, \delta_{ij}, \rho - \tau\right) \tag{3}$$

$$\text{s.t.} \frac{\delta_{ij}}{K} \sum_{k=1}^{K} d_{\mathbf{W}_k}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right) \leq \delta_{ij} \tag{4}$$

where $H(\mathbf{W}_k,\mathbf{W}_0)$ is LogDet divergence [3], $d_{\mathbf{W}_k}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right) = \left(\mathbf{B}_i^k, \mathbf{B}_j^k\right)^T \mathbf{W}_k \left(\mathbf{B}_i^k, \mathbf{B}_j^k\right)$, n is the number of training samples, γ is a tradeoff parameter, the initial matrix $\mathbf{W}_0$ is an identity matrix. ρ is threshold and τ is a neighborhood parameter. δ_{ij}=1 if the two samples belong to the same person, otherwise, δ_{ij}=−1. Thus, the distance between two samples should be less than $\rho-\tau$ if the two samples belong to the same person, otherwise $\rho+\tau$. The first order continuous differentiable hinge loss function $l(\cdot,\cdot)$ is defined as follows.

$$l(x, x_o) = \begin{cases} 0 & x \leq x_o \\ (x, x_o)^2 & x > x_o \end{cases} \tag{5}$$

Fig. 3 All images of the first people

Fig. 4 The misaligned face images of Fig. 3

To compute the distance using the IPMML method, the final features of a testing sample of the 1st person are given in Table 2, denoted as T11. The feature vector is divided to 7 blocks with size of 1×2 too. First, the distance between the 1st testing samples and every training sample S*ij* (S*i* is listed in the row, and *j* is in the column) used the IPMML method is computed with the threshold ρ=1×10^5 according to the reference [20], as shown in Table 3. To compare with the distance using the PMML method and the Euclidean distance, the PMML distance and Euclidean distance between the 1st testing samples and every training sample are computed, as shown in Tables 4 and 5. It can be observed that three distances can be identified well using the NNC method, but the distance difference using the IPMML method (20 times) is more obvious than the ones using the PMML distance (10 times) and Euclidean distance (2 times). Therefore, the IPMML method is used to compute the distance in this paper.

2.4 The realization of the proposed face recognition method

According to the above analysis, the detailed realization process is as follows.

- Step 1: LBP operator is used to compute the LBP values of training samples $\mathbf{A}_i$ (i=1,⋯,N) and then the LBP images are obtained.
- Step 2: A set of eigenvectors is got using LDA and then the LBP images are projected on the eigenvectors to obtain the final features$\mathbf{B}_i$ with low dimension.
- Step 3: Training image blocks $\mathbf{B}_i^k$, (k=1,⋯,K) is obtained according to the dimension of $\mathbf{B}_i$ and every block is changed to a column vector.
- Step 4: Testing image blocks $\mathbf{B}_j^k$, (k=1,⋯,K) is obtained according to step 1~3.
- Step 5: The discriminative distance $d(\mathbf{B}_j,\mathbf{B}_i)$ is calculated between training image blocks $\mathbf{B}_i^k$ and testing image blocks $\mathbf{B}_j^k$ according to the algorithm as follows [2].
- Step 5.1: Set t=1, $\mathbf{W}_k^1 = \mathbf{W}_0$, η_{ij}=0, $\xi_{ij}^t = \delta_{ij}\rho - \tau$, ρ=1×10^4, τ=1.0, γ=0.1.

Table 6 The ARAs of original Yale dataset (%)

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	82.72	92.59	88.15	89.88	90.62	90.86	94.57	86.17	72.13	97.08	78.17
3	88.33	93.61	94.72	92.22	96.67	96.67	97.78	92.29	87.20	96.95	82.03
4	94.60	97.78	97.14	97.78	98.10	98.10	99.05	91.44	83.95	99.17	85.83
5	95.19	98.52	98.89	98.15	99.26	99.26	100.00	92.27	87.26	99.04	92.05
6	96.00	98.22	97.78	97.33	99.11	99.11	99.56	94.33	88.70	99.06	91.22
7	98.89	98.89	99.44	98.33	100.00	100.00	100.00	96.89	96.79	98.92	95.05

Table 7 The SRs of original Yale dataset

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	3.50	5.59	8.25	7.12	6.72	6.30	2.60	5.53	14.10	0.59	6.41
3	7.41	2.41	0.96	2.93	3.00	3.00	2.55	3.71	4.42	3.19	12.60
4	2.91	3.06	1.90	3.06	0.95	0.95	0.95	6.29	5.33	0.79	4.75
5	2.80	0.64	0.00	1.28	1.28	0.64	0.00	6.47	9.48	1.03	4.34
6	3.53	1.54	2.04	0.00	1.54	1.54	0.77	4.79	7.11	0.58	2.60
7	2.89	0.96	1.67	4.41	0.00	7.64	0.00	2.81	3.35	0.66	2.27

- Step 5.2: Compute the distance $d_{\mathbf{W}_k^t}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right)$ between the k-th block of testing sample $\mathbf{B}_j^k$ and the k-th block of training sample $\mathbf{B}_i^k$, (k=1,⋯,K).
- Step 5.3: Solve α according to eq. (6) and set α=min(α,η_{ij}), η_{ij}=η_{ij}−α.

$$\frac{\delta_{ij}}{K}\sum_{k=1}^{K}\frac{d_{\mathbf{W}_k^t}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right)}{1-\delta_{ij}ad_{\mathbf{W}_k^t}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right)}-\left(\xi_{ij}^t\frac{n}{2\gamma}\alpha\right)=0 \tag{6}$$

– Step 5.4: Update $\mathbf{W}_k^{t+1}$ according to eq. (7) and set $\mathbf{W}_k = \mathbf{W}_k^{t+1}$, ($k$=1,⋯,$K$),

$$\mathbf{W}_k^{t+1} = \mathbf{W}_k^t + \mu\left(\mathbf{W}_k^t\left(\mathbf{B}_i^k-\mathbf{B}_j^k\right)\left(\mathbf{B}_i^k-\mathbf{B}_j^k\right)^T\mathbf{W}_k^t\right) \tag{7}$$

where, $\mu = \delta_{ij}\alpha/\left(1-\delta_{ij}ad_{\mathbf{W}_k^t}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right)\right)$.

– Step 5.5: Update ξ_{ij}^{t+1} according to eq. (8).

$$\xi_{ij}^{t+1} = \xi_{ij}^t-\frac{n}{2\gamma}\alpha \tag{8}$$

Table 8 The ARAs of misalignment Yale dataset (%)

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	62.96	70.62	59.51	65.19	61.98	62.72	71.85	60.34	47.78	70.25	40.50
3	72.50	77.78	69.17	74.17	75.00	75.28	83.06	71.14	59.15	82.16	55.62
4	83.17	91.75	78.10	90.48	83.49	84.44	90.79	72.67	62.47	88.94	69.56
5	85.56	91.11	83.70	89.26	86.30	86.30	94.07	78.00	70.22	92.21	68.69
6	88.44	88.89	83.11	88.00	90.22	89.78	94.22	84.67	74.45	91.22	72.67
7	85.56	88.33	78.89	86.67	77.22	76.67	88.33	79.78	75.06	86.78	73.43

Table 9 The SRs of misalignment Yale dataset

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	13.17	10.59	6.88	12.24	7.92	8.19	12.19	8.35	15.52	27.99	18.94
3	4.33	13.75	8.21	15.57	9.61	10.55	12.20	11.66	11.12	4.99	9.42
4	3.85	5.25	9.38	4.76	5.42	6.12	3.85	10.98	6.89	4.09	4.93
5	5.09	2.94	1.70	4.63	7.56	7.56	4.63	11.73	7.75	3.89	5.26
6	6.01	1.54	2.04	1.33	2.78	3.36	2.04	8.16	15.07	9.10	12.40
7	14.56	9.62	6.01	0.00	10.14	3.47	9.62	14.42	10.93	13.08	9.09

– Step 5.6: Stop the iteration and output $\mathbf{W}_k$ if $\frac{\delta_{ij}}{K}\sum_{k=1}^{K} d_{\mathbf{W}_k}\left(\mathbf{B}_j^k, \mathbf{B}_i^k\right) \le \xi_{ij}$ or $t>(N+M)\times 10$, otherwise, $t=t+1$ and go to Step 3.
– Step 5.7: Compute the distance between $\mathbf{B}_j$ and $\mathbf{B}_i$ according to eq. (2) using the optimal Mahalanobis matrix $\mathbf{W}(\mathbf{W}_1,\mathbf{W}_2,\cdots,\mathbf{W}_k)$.
– Step 6: If $d(\mathbf{B}_j,\mathbf{B}_i) = \min d_i(\mathbf{B}_j,\mathbf{B}_i)$ and $\mathbf{B}_l \in c (1\le c\le C)$, then $\mathbf{B}_j \in c$, where, C is the number of training samples categories.

3 Experimental results and analysis

In this paper, we do experiments on Yale, Extended Yale B (EYB) and AR datasets [7]. Yale dataset includes 165 images of 15 individuals and each of them has 11 images under different illuminations and expressions. EYB dataset contains 16,128 images of 38 people in total and 10 people who have 64 images under different lightings are selected in this paper. AR dataset includes 3276 images of 126 people and each one has 26 images under various lighting, facial expressions and facial details. In addition, to observe the recognition performance under misalignment conditions, the first 10 pixels of the top left corner are removed and we conduct this on the first 3 images of each one, so the misaligned face images are obtained. The whole images of the first people are shown in Fig. 3. The corresponding misaligned face images are shown in Fig. 4.

3.1 The yale dataset

2~7 images of each person are randomly selected as the training samples for both original and misalignment Yale dataset. First, LBP operator is used to extract texture features. Then, LDA is used to decrease the dimension of LBP images. Third, the final features are divided into 7 blocks and the size of every block is 1×2 since the dimension of the final features is 14. Fourth,

Table 10 The ARAs of original EYB dataset (%)

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	55.81	95.00	90.54	93.87	91.29	91.51	96.13	72.26	62.82	91.36	70.57
3	77.76	96.50	98.14	95.90	99.23	99.29	99.95	77.60	82.37	95.91	77.71
4	80.56	99.33	99.06	99.22	99.22	99.17	99.83	90.49	88.51	95.74	80.32
5	89.72	99.66	99.60	99.49	99.72	99.72	100.00	94.48	85.79	96.71	86.37
6	87.93	99.60	99.37	99.54	99.48	99.43	99.60	95.19	91.56	98.51	91.22
7	89.36	99.65	99.53	99.65	99.82	99.82	99.94	94.52	96.43	98.29	93.32

Table 11 The SRs of original EYB dataset

*n**	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	18.67	5.46	6.83	6.02	8.08	8.02	4.73	20.84	11.28	7.31	10.72
3	6.80	4.25	1.05	5.13	0.34	0.25	0.09	11.89	7.24	1.98	4.36
4	5.95	0.33	0.25	0.19	0.35	0.33	0.17	4.62	6.99	3.86	3.33
5	4.26	0.17	0.26	0.34	0.10	0.10	0.00	4.36	6.53	2.03	5.78
6	1.64	0.10	0.53	0.20	0.46	0.43	0.43	4.08	5.95	0.63	3.53
7	2.81	0.00	0.10	0.00	0.18	0.18	0.10	8.83	2.36	1.31	3.35

NNC method is used to classify face images according to the Euclidean distance and discriminative distance based on PMML and IPMML. For comparison, the face recognition methods used LDA+IPMML [20], LBP+IPMML, LBP+PMML, LBP+NNC,, LBP+LDA+PMML, LBP+LDA+NNC, LBP+LDA+IPMML(the proposed method), Convolutional Neural Networks (CNN), LBP+CNN, CosFace [5] and LBP+CosFacemethods are performed. For convenience to present, they are denoted as LIP, LBIP, LBPP, LBN, LLPP, LLN, LLIP, CNN, LC, CF and LCF in the following tables. The structure of the CNN includes two convolutional layers with 4 and 8 kernels (3×3) and two pooling layers (2×2). The training parameters are set: the learning rate is 0.001. The learning rate decay is 1E-6, momentum is 0.9, the epoch is 120, the batch size is 20. The network in CF and LCF is fine-tuned using the training samples, a sample is randomly selected from a person as the validation sample, and the rest are taken as the test samples in the CosFace method. The methods are performed on original and misalignment images. Every experiment is performed 5 times and the Average Recognition Rate (ARA) and the standard deviation (SR) are calculated. The results are shown in Tables 6, 7, 8 and 9. It is obviously that the proposed method is better than other methods for face images with different illumination and expression, especially the number of training samples is small.

Table 12 The ARAs of misalignment EYB dataset (%)

*n**	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	51.56	91.34	86.34	89.68	87.04	87.26	94.41	55.26	60.18	74.33	62.15
3	64.97	92.84	93.61	91.58	94.81	94.86	97.81	60.35	73.43	78.58	69.93
4	70.17	95.17	94.22	94.89	94.50	94.50	97.67	74.96	83.60	86.12	74.71
5	72.82	96.95	95.93	96.16	95.76	95.76	98.42	82.26	81.69	83.66	86.21
6	81.67	95.80	94.37	94.94	94.43	94.37	97.41	83.42	88.71	85.64	84.66
7	83.39	98.42	97.72	98.07	97.72	97.72	98.83	84.39	93.14	92.82	92.30

Table 13 The SRs of misalignment EYB dataset

*n**	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	14.25	5.37	7.08	6.18	7.97	8.05	4.53	21.41	10.93	7.60	15.80
3	12.89	4.18	1.18	5.08	0.09	0.09	0.62	17.09	11.57	15.06	12.92
4	9.82	0.33	0.19	0.25	0.33	0.33	0.33	11.71	6.66	3.24	8.39
5	7.12	2.00	2.20	1.93	1.63	1.63	0.64	11.28	7.03	6.15	6.02
6	2.85	0.40	0.36	0.40	0.61	0.61	0.69	6.93	3.85	5.37	4.11
7	6.34	1.98	3.04	2.44	2.90	2.90	1.72	9.02	2.22	6.01	1.55

Table 14 The ARAs of original AR dataset (%)

*n**	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	FC	LCF
2	63.99	74.90	76.94	72.22	81.88	81.98	88.26	24.83	39.35	79.57	48.33
5	82.90	86.55	90.52	84.33	92.38	92.18	95.75	58.69	68.26	93.83	66.38
7	96.58	89.25	94.69	87.76	96.27	96.36	97.94	81.60	83.11	93.50	73.71

3.2 The EYB dataset

We randomly choose 2~7 images of each person as training samples for both original and misalignment EYB dataset. First, LBP operator is used to extract texture features. Second, LDA is used decrease the dimension of LBP images. Third, the final features are divided into 3 blocks and the size of every block is 1×3 since the dimension of the final features is 9. Fourth, NNC method is used to classify face images according to the Euclidean distance and discriminative distance based on PMML and IPMML. The same methods with section 4.1 are performed on original and misalignment images. Every experiment runs 5 times and the ARA and SR are calculated. The results are shown in Tables 10, 11, 12 and 13. It is obviously that the proposed method is better than other methods for face images with different illumination conditions, especially the number of training samples is small.

3.3 The AR dataset

We choose 2, 5, 7 images of each person as training samples for both original and misalignment AR dataset. First, LBP operator is used to extract texture features. Second, LDA is used to decrease the dimension of LBP images. Third, the low dimensional images are divided into 13 blocks and the size of every block is 1×3 since the dimension of the final feature is 39. Fourth, NNC method is used to classify face images according to the Euclidean distance and discriminative distance based on PMML and IPMML. The same experiments are performed like section 4.1 and 4.2. The similar results are obtained, as shown in Tables 14, 15, 16 and 17. They prove that the proposed method is effective and robust for different illumination and facial expression.

Table 15 The SRs of original AR dataset

*n**	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	FC	LCF
2	2.99	7.47	7.29	7.18	5.23	5.23	4.46	24.06	11.07	10.17	13.32
5	10.05	2.11	1.70	1.96	1.56	1.82	1.60	24.36	12.88	0.99	3.01
7	1.39	4.34	2.22	4.68	1.54	1.44	0.77	6.95	9.70	2.35	3.36

Table 16 The ARAs of misalignment AR dataset (%)

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	53.16	65.45	66.42	62.85	70.76	70.87	77.47	21.88	32.88	64.59	33.05
5	77.02	83.49	74.92	80.36	83.41	83.33	88.77	54.33	54.68	82.81	45.95
7	84.43	80.00	81.58	77.98	84.17	84.17	87.19	68.57	63.60	79.37	57.91

Table 17 The SRs of misalignment AR dataset

n*	LIP	LBIP	LBPP	LBN	LLPP	LLN	LLIP	CNN	LC	CF	LCF
2	2.68	7.28	7.01	6.97	5.18	5.18	4.66	17.60	10.74	4.30	7.49
5	6.25	5.12	6.82	5.54	4.20	4.09	3.07	14.04	16.73	2.70	10.67
7	2.43	3.62	2.34	3.99	4.80	4.80	5.05	9.17	16.64	3.21	4.65

From Table 4~17, it can be found that the proposed LLIP method can achieve higher recognition rate and less standard deviation for both original and misalignment images than other methods, which illustrates that the features extracted using LBP+LDA has more discrimination information and IPMML is better than PMML and NNC.

4 Conclusions

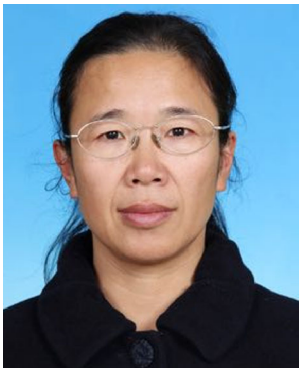
In this paper, a face recognition method based on LBP and IPMML is proposed. First, LBP operator is used to extract texture features. Then, LDA is used to decrease feature dimension. Third, IPMML is used to compute the discriminative distance. Finally, NNC is applied to classify face images. The features extracted using LBP+LDA method have rich texture information and low dimension, which reduces the complexity of calculation. Additionally, IPMML is also robust for misalignment images. Experimental results show that the proposed method can recognize face images reliably and effectively for both original and misalignment images.

Acknowledgements This work was supported by Shandong Province Natural Science Foundation, China: ZR2016FQ14, the Key Research and Development Programs of Shandong Province Project under grant 2018GGX101040, Science and Technology Research Program for Colleges and Universities in Shandong Province under Grant J18KA315.

References

1. Belhumeur PN, Hespanha JP, Kriengman DJ (1997) Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19:711–720
2. Cui Z, Li W, Xu D et al (2013) Fusing Robust Face Region Descriptors via Multiple Metric Learning for Face Recognition in the Wild. IEEE Conference on Computer Vision & Pattern Recognition:3554–3561
3. Davis J V, Kulis B, Jain P, et al (2007) Information-theoretic metric learning. Proceedings of the 24th international conference on Machine learning. ACM, pp. 209–216.
4. Hao S, Wei W, Yan Y, Lorenzo B (2017) Class-wise dictionary learning for hyperspectral image classification. Neurocomputing 220:121–129
5. He L, Li H, Zhang Q, Sun Z (2018) Dynamic Feature Learning for Partial Face Recognition. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 7054–7063
6. Ke P, Cai M, Wang H, Chen J. (2018) A novel face recognition algorithm based on the combination of LBP and CNN. 14th IEEE International Conference on Signal Processing Proceedings, pp. 539–543
7. Lee KC, Ho J, Kriegman DJ (2005) Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans Pattern Anal Mach Intell 27(5):684–698
8. Ojala T, Pietikainen M, Harwood D (1996) A comparative study of texture measures with classification based on feature distributions. Pattern Recogn 29(1):51–59
9. Wang J, Li Z (2018) Research on Face Recognition Based on CNN. 2nd International Symposium on Resource Exploration and Environmental Science, pp. 1–5

10. Wang H, Wang Y , Zhou Z, et al (2018) CosFace: Large Margin Cosine Loss for Deep Face Recognition. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5265–5274
11. Wang W, Yan Y, Stefan W, Nicu S (2016) Category specific dictionary learning for attribute specific feature selection. IEEE Trans Image Process 25(3):1465–1478
12. Wang W, Yan Y, Zhang L, Hong R, Nicu S (2016) Collaborative sparse coding for multiview action recognition. IEEE MultiMedia 23(4):80–87
13. WANG X, ZHANG Y, MU X, ZHANG F (2012) The Face Recognition Algorithm Based on Improved LBP. Opto-Electronic Engineering 39(7):109–114
14. W. Wang, C. Zhen, Y. Yan, J. Feng, S. Yan, X. Shu, and S. Nicu (2016) Recurrent face aging. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2378–2386
15. Wen L, Xi C, LIU Z, HUANG Q (2015) Face recognition based on self-adaptive fusion of LBP and CS-LDP. Journal of Shanxi Normal University (Natural Science Edition) 43(4):48–53
16. Yu H, Yang J (2001) A Direct LDA Algorithm for High-dimensional Data with Application to Face Recognition. Pattern Recogn:2067–2070
17. Yulin W, Mingyan J (2018) Face recognition system based on CNN and LBP features for classifier optimization and fusion. The Journal of China Universities of Posts and Telecommunications 25(1):37–47
18. Zhang H, Qu Z, Yuan L, et al (2017) A face recognition method based on LBP feature for CNN. 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). IEEE
19. Zhao Y, Gao J, Wang R, Hu J (2004) An Extended Nearest Neighbor Method Based on Bionic Pattern Recognition. *Acta Electron Sin* 12:196–198
20. Zhou L, Wang H, Lu Z, Nie T, Zhao K (2016) Face Recognition Based on LDA and Improved Pairwise-Constrained Multiple Metric Learning Method. Journal of Information Hiding and Multimedia Signal Processing 7(5):1092–1099
21. ZHU C, DING Y, YUAN B, CAO H (2015) Face Recognition Based on Local Binary Pattern and Local Phase Quantization. Journal of Nanjing Normal University (Natural Science Edition) 38(1):104–107



Lijian Zhou received the B.S. degree and the M.S. degree from Daqing Petroleum Institute in 1993 and 1996, and the Ph.D. degree in 2007 from Ocean University of China. She is currently in the schoolof Information and Control Engineering, Qingdao University of Technology. Her research interests include image processing, pattern recognition, and intellectual property protection.

Hui Wang received the B.S. degree and the M.S. degree from Qingdao University of Technology in 2012 and 2015. She has been with Jinan Technican College. Her research interests include image processing and pattern recognition.

Shanshan Lin received the B.S. degree from Huai hai Institute of Technology in 2017. Currently she is studying at Qingdao University of Technology, Information and Control Engineering. Her research interests include image processing and pattern recognition.

Siyuan Hao received her PhD degree from the College of Information and Communications Engineering in Harbin Engineering University, Harbin, China, in 2015. She is currently a Researcher at Qingdao University of Technology, China, where she teaches remote sensing and electrical communication. Her research interests focus on hyperspectral imagery processing and machine learning.

Zhe-Ming Lu received the B.S. and M.S. degrees in Electrical Engineering from Harbin Institute of Technology (HIT), Harbin, P. R. China, in 1995 and 1997 respectively, and the Ph. D. degree in instrument science and technology from HIT, Harbin, P. R. China, in 2001. He is currently working as a full professor in Zhejiang University, Hangzhou, China. His research interests include multimedia single analysis and processing, information hiding and astronautics signal processing, etc.

Affiliations

Lijian Zhou[1] · Hui Wang[1,2] · Shanshan Lin[1] · Siyuan Hao[1] · Zhe-Ming Lu[3]

[1] School of Information and Control Engineering, Qingdao University of Technology, Qingdao 2660520, People's Republic of China

[2] Jinan Technician College, Jinan 2660520, People's Republic of China

[3] School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, People's Republic of China