International Journal of Security and Its Applications
Vol.8, No.5 (2014), pp.253-264
http://dx.doi.org/10.14257/ijsia.2014.8.5.23

A Block Discriminant Analysis for Face Recognition

Peng Cui [1]

School of Computer Science & technology, Harbin University of Science and Technology, Harbin, China
E-mail:cuipeng83@163.com

Abstract

As for illumination variation, traditional feature extraction methods are not satisfactory for face recognition. A block discriminant analysis algorithm is proposed to solve the problem. Firstly, local contrast enhancement is used to compensate for uneven illumination; secondly, discrete cosine transform (DCT) is implemented for divided image blocks. According to data distribution of DCT matrix, the block candidate features are selected, and merged to candidate features; finally, block discriminant analysis are carried out for features extraction. Experiments are tested on Yale and Yale B, the results prove the algorithm outperform related algorithms.

Keywords: *Face recognition, Feature extraction, Discrete cosine transform, Discriminant analysis*

1. Introduction

It is difficult to recognize face under illumination variation. If the uneven illumination is not processed, recognition rate cannot be satisfactory. The performance of a face recognition system mainly depends on methods of feature extraction and classification. At present, many algorithms of feature extraction need to reduce the dimensionality of images. These algorithms are divided into two categories: spatio domain and spatio-frequency domain [1-5]. If PCA is used for feature extraction, the images in database of spatio domain usually are used as input. LDA can search nearest distances from input vectors in same classes, and the farthest distances from input vectors in different classes. However, methods based on LDA often are affected by small samples problem, eigen value cannot be calculated. Many works modify LDA to solve small samples problem. Traditional methods such as D-DCT [1], Eigenface[6], Fisherface[7], DLDA[8] are all based on global feature extraction, so they are sensitive to illumination. As a result, local details can be ignored.

Some algorithms that use features in spatio-frequency domain can obtain advantages of dimension reduction. For example, Discrete Cosine Transform (DCT) and Wavelet Transform (WT) select low frequency coefficients as eigen value for classification. These algorithms can not only reduce dimensionality, but also obtain robustness on illumination variation by data transform and coefficients selection.

2. Local Contrast Enhancement for Preprocessing Images

Most of face recognition systems use image enhancement measures for overcoming illumination variation, which retain visual effect of original scenes and minimize

[1]Corresponding author. Tel.: +86 45157660019.
E-mail *address*: cuipeng83@163.com.

ISSN: 1738-9976 IJSIA

distortion. Two methods are used for contrast enhancement: global Histogram Equalization (HE) and Local Contrast Enhancement (LCE). HE methods have low computation complexity, and hold whole illuminance. However, compressed data ranges in global view can lead to lose some tiny textures. Because using histogram information of entire image can constrain contrast stretching ratio. LCE can effectively enhance visuality of tiny textures [9, 10]. Although it leads to visual gradient reverse or cause undesirable halo, the problems can be solved by feature selection.

Image parameters under uneven light source are unstable. Local contrast enhancement (LCE) is useful for image details. LCE can heighten unobserved but important texture. As LCE is not sensitive to illumination variation, image contrast is greatly enhanced. Furthermore, image texture can be steadily extracted.

We define a square or rectangle neighborhood, and move the center of neighborhood from one pixel to another. $\overline{I(m,n)}$ stands for average luminance of (m,n) pixel's neighborhood whose size is $l \times l$. For a $M \times N$ image, $\overline{I(m,n)}$ can be represented as in equation (1)

$$\overline{I(m,n)} = \frac{1}{l \times l} \sum_{i=-(l-1)/2}^{(l-1)/2} \sum_{j=-(l-1)/2}^{(l-1)/2} I(m+i, n+j) \tag{1}$$

as luminance of pixel in position (m,n) is $I(m,n)$, so in equation (2), its local contrast can be represented as in equation (2)

$$e_{mn} = \begin{cases} \log\left(I(m,n) / \overline{I(m,n)}\right) & \text{if } I(m,n) > T \text{ and } \overline{I(m,n)} > T \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

T is threshold value, which is initialized by 1; e_{mn} represents value of the image E in position (m,n).

Local contrast is not the original luminance, but the ratio of pixel luminance to neighborhood average. Hence, range of image data is compressed. As local contrast can be positive or negative, standardization are implemented as in equation (3)

$$\boldsymbol{E}_{normalized} = (\boldsymbol{E} - \min(\boldsymbol{E}(:)) / (\boldsymbol{E} - \max(\boldsymbol{E}(:)) \tag{3}$$

$\boldsymbol{E}_{normalized}$ represents standardization matrix of local contrast; min (E(:)) means minimal value among all local contrast and max(E(:)) means maximal value among all local contrast.

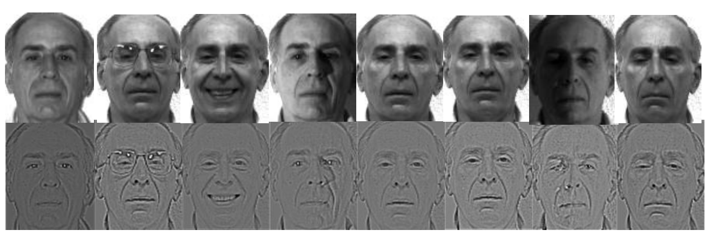

Figure 1. Local Contrast Enhancement Images from Yale with 5×5 Pixel Neighbor Window

In Figure 1, images from Yale database are processed by LCE. The first row images are original images from Yale database, the second row images are enhanced images by LCE with 5×5 pixel neighbor window. We can see the smaller neighborhoods are selected, the better details are enhanced.

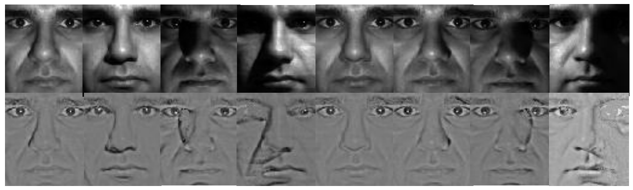

Figure 2. Local Contrast Enhancement Images from Yale B with 7×7 Pixel Neighbor Windows

In Figure 2, images from Yale B database are also processed by LCE. The first row images are original images from Yale B database, the second row images are enhanced images by LCE with 7×7 pixel neighbor window. The problem of illumination variation is effectively weaken. Most textures of images are retained and enhanced.

3. Block Discriminant Analysis for Feature Extraction

3.1. Block DCT Transform

DCT has been successfully used for face recognition [11-14]. We divide enhanced images into blocks, and carry out DCT to get spatio-frequency coefficient for each. If image size is $M \times N$ and block size is $b \times b$, let $b=8$, $b_r=M/8$, $b_c=N/8$, then get $b_r \times b_c$ blocks. In equation (4), image $\boldsymbol{E}$ is represented as

$$\boldsymbol{E} = \begin{bmatrix} \boldsymbol{E}_{1,1} & \boldsymbol{E}_{1,2} & \cdots & \boldsymbol{E}_{1,b_c} \\ \boldsymbol{E}_{2,1} & \boldsymbol{E}_{2,2} & \cdots & \boldsymbol{E}_{2,b_c} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{E}_{b_r,1} & \boldsymbol{E}_{b_r,2} & \cdots & \boldsymbol{E}_{b_r,b_c} \end{bmatrix}_{b_r \times b_c} \tag{4}$$

DCT is processed for per block, so 64 DCT coefficients are obtained, which cover all sptio-frequency component of image. 2D-DCT coefficients of an 8×8 block are represented as $\boldsymbol{D}_{ij}$ in equation (5)

$$\boldsymbol{D}_{i,j} = \boldsymbol{H}\boldsymbol{E}_{ij}\boldsymbol{H}^T, \quad 1 \le i, 1 \le j \tag{5}$$

which $\boldsymbol{H}$ is 8×8 matrix, represented as Table 1.

Table 1. 8×8 Transform Matrix *H* for Image Block

Row	Column							
	1	2	3	4	5	6	7	8
1	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536	0.3536
2	0.4904	0.4157	0.2778	0.0975	-0.0975	-0.2778	-0.4157	-0.4904
3	0.4916	0.1913	-0.1913	-0.4619	-0.4619	-0.1913	0.1913	0.4619
4	0.4157	0.0975	-0.4904	-0.2778	0.2778	0.4904	0.0975	-0.4157
5	0.3536	0.3536	-0.3536	0.3536	0.3536	-0.3536	-0.3536	0.3536
6	0.2778	0.4904	0.0975	0.4157	0.4157	-0.0975	0.4904	-0.2778
7	0.1913	0.4619	0.4619	-0.1913	-0.1913	0.4619	-0.4619	0.1913
8	0.0975	0.2778	0.4157	-0.4904	-0.4904	-0.4157	0.2778	-0.0975

Definition 1. let data $\boldsymbol{x}$ represents a image, $\boldsymbol{x} \in \boldsymbol{R}^{m \times n}$, and divides it into $(m/l)*(n/l)$ blocks, which each block size is $l \times l$. After DCT, select top k coefficient as principal features, k<<l*l, these features are called block candidate features.

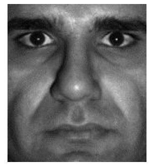

Figure 3. A image from Yale B Database

Figure 3 shows an image from Yale B database. Figure 4 is block DCT matrix of selected image. With rectangular are as represent the DCT coefficient matrix of each image block. Within 8×8 blocks, frequency coefficient still conforms to the distribution of DCT coefficients. Larger DCT coefficients are mainly in DC component and low frequency AC components. DC component and top 3 AC components are selected as candidate features, so number of candidate features is $b_r \times b_c \times 4$.

	1	2	3	4	5	6	7	8	9	10
1	492.13	-13.97	-0.49	2.25	-3.37	-0.23	-0.20	-0.29	574.13	-44.20
2	1.83	2.54	-0.43	-3.81	2.87	0.10	-0.03	0.07	-11.08	7.58
3	1.51	1.25	-6.27	2.95	0.20	0.19	-0.39	0.15	-6.80	4.12
4	-2.15	-4.10	3.74	2.59	-0.35	0.37	0.25	0.37	0.16	-2.50
5	-3.62	-4.49	0.49	0.12	-0.13	0.05	0.20	0.16	1.63	0.02
6	2.67	-0.14	-0.04	0.19	-0.17	0.09	-0.42	-0.59	0.12	0.20
7	0.05	-0.36	-0.14	0.33	-0.49	-0.05	0.27	0.28	0.25	0.22
8	-9.49	-0.11	0.12	-0.12	0.26	-0.04	0.47	-0.22	0.33	-0.35
9	484.25	-21.11	-6.92	-2.19	-3.00	-0.22	-0.38	0.10	549.63	-33.91
10	0.41	4.12	-0.42	2.52	3.43	0.25	0.05	-0.24	29.04	-0.12
11	5.85	2.22	-5.78	-0.04	0.33	-0.06	0.12	0.61	-21.34	-5.28
12	1.14	0.16	0.36	-3.21	-0.29	0.46	0.10	0.14	6.92	3.88
13	-4.25	-6.02	-4.50	0.33	0.00	0.04	0.24	-0.27	-2.12	0.24
14	0.26	-0.15	-0.54	8.27	0.14	-0.01	-0.49	-0.02	3.21	4.30
15	-0.26	-0.08	-0.12	-0.03	0.14	0.06	-0.47	-0.42	-0.23	-0.07
16	8.67	0.07	-0.28	-0.07	-0.11	0.03	-0.27	0.09	9.27	-0.45
17	274.25	56.29	1.67	-6.19	-3.50	0.22	0.12	0.11	212.13	-1.06
18	106.69	18.05	-19.98	-6.16	-6.23	-0.15	-0.28	0.15	58.08	-22.96

Figure 4. DCT Matrix Based on 8×8 Blocks of Selected Image

3.2 Calculation for Block Discriminant Factor

The traditional feature extraction methods distinguish different face by eigen value and eigenvector. To solve small sample problem, the dimensionality of images are need to be reduced. The proposed Block Discriminant Analysis (BDA) reduces the data dimension by calculating Block Discriminant Factor (BDF), and needn't calculate eigen values.

Definition 2. let $\tilde{\boldsymbol{C}}$ be candidate feature matrix of P training samples which belong to c different classes, which column vector of $\tilde{\boldsymbol{C}}$ is candidate feature. Then $\tilde{\boldsymbol{C}}_k$ represents sample set which belong kth class. Let BDF be the ratio of between-class variance to total variance.

The calculation steps are as follows:

- Let $K=b_r*b_c*4$. After DCT, the candidate feature matrix of P images can be represented as

$$\tilde{\boldsymbol{C}}=\begin{bmatrix}\tilde{c}_{1,1} & \tilde{c}_{1,2} & \cdots & \tilde{c}_{1,P}\\ \tilde{c}_{2,1} & \tilde{c}_{2,2} & \cdots & \tilde{c}_{2,P}\\ \vdots & \vdots & \vdots & \vdots\\ \tilde{c}_{K,1} & \tilde{c}_{K,2} & \cdots & \tilde{c}_{K,P}\end{bmatrix}_{K\times P} \tag{6}$$

$\tilde{\boldsymbol{C}}$ is candidate feature matrix, which is enhanced by LCE and reduced dimensionality by DCT. Column vector of $\tilde{\boldsymbol{C}}$ is candidate feature vector of corresponding image, row vector of $\tilde{\boldsymbol{C}}$ is same candidate feature of different images, $\tilde{c}_{i,j}$ means ith candidate feature of jth image, $1\le i\le K$, $1\le j\le P$. Figure 5 and Figure 6 illustrate the enhanced images from Yale and Yale B by LCE respectively, which block-size is 8×8 and each image is divided into 64 blocks, and reconstructed images with top 4 important DCT coefficients. The first row images are enhanced images, and the second row images are reconstructed images.

Figure 5. Reconstructed Images from Yale by LCE with 4 Features per Block

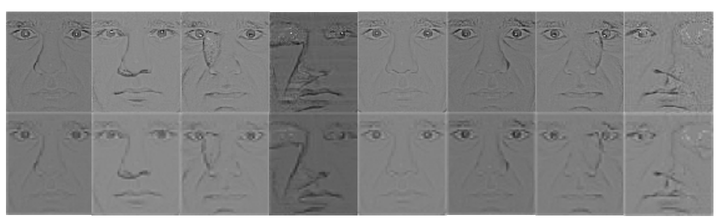

Figure 6. Reconstructed Images from Yale B by LCE with 4 Features per Block

- According to labeled data set, between-class variance of candidate features is represented as

$$\tilde{\delta}_{ij}^{B}=\sum_{k=1}^{c}P_k\left(\frac{1}{P_k}\sum_{\tilde{c}_{ij}\in\tilde{C}_k}\tilde{c}_{ij}-\frac{1}{P}\sum_{u=1}^{P}\tilde{c}_{iu}\right)^2 \tag{7}$$

P_k is number of samples which belongs to kth class; $\frac{1}{P_k}\sum_{\tilde{c}_{ij}\in\tilde{C}_k}\tilde{c}_{ij}$ is mean value of candidate features which belongs to kth class in position (i, j); $\tilde{\delta}_{ij}^{B}$ is the between-class variance of jth image of ith feature; $\frac{1}{P}\sum_{u=1}^{P}\tilde{c}_{iu}$ is mean value of ith candidate feature of all samples； $\tilde{c}_{ij}\in\tilde{C}_k$ means the candidate features which belong to kth class images.

- Calculate variance of candidate features is represented as in ：

$$\tilde{\delta}_{ij}^{T} = \sum_{v=1}^{P}\left(\tilde{c}_{iv} - \frac{1}{P}\sum_{u=1}^{P}\tilde{c}_{iu}\right)^{2} \tag{8}$$

which $\frac{1}{P}\sum_{u=1}^{P}\tilde{c}_{ij}(u)$ is mean value of candidate feature of all samples which $\tilde{c}_{ij}$ are in position (i, j). $\tilde{\delta}_{ij}^{T}$ means variance of ith candidate feature of jth image.

- Calculate BDF of position (i, j) ：

$$\alpha_{ij} = \frac{\tilde{\delta}_{ij}^{B}}{\tilde{\delta}_{ij}^{T}}, \qquad 1 \le i \le K, 1 \le j \le P \tag{9}$$

which α_{ij} represents BDF value of α in position (i, j).

- Standardize BDF ：

$$\boldsymbol{\alpha}_{normalized} = \frac{\left(\boldsymbol{\alpha} - \min\left(\boldsymbol{\alpha}(:)\right)\right)}{\left(\max\left(\boldsymbol{\alpha}(:)\right) - \min\left(\boldsymbol{\alpha}(:)\right)\right)} \tag{10}$$

which $\min(\boldsymbol{\alpha}(:))$ means minimum value of $\boldsymbol{\alpha}$, $\max(\boldsymbol{\alpha}(:))$ means maximum value of $\boldsymbol{\alpha}$. Low frequency and DC components correspond to the larger values of $\boldsymbol{\alpha}$, while mid frequency coefficients correspond to the smaller values.

3.3. LCE+BDA Algorithm

Like existing LDA and PCA, BDA is also a statistical method. Its validity is greatly affected by the number of training samples. When each class has enough samples available, BDA is a optimized DCT feature extraction method.

The process of LCE+BDA algorithm is as follows:

- LCE is carried out for all training samples. To select a small neighborhood, in terms of equation (1−3) calculate the ratio of per pixel to mean value of its neighborhood, and obtain local contrast and standardize it.
- In terms of equation (4-5), calculate block DCT matrix. In terms of mask, select important coefficients as block candidate features, then combine candidate features of all blocks as candidate feature vector of images, and obtain candidate feature matrix of training image set.
- In terms of equation, (7-8), calculate between-class variance and total variance of candidate features of images.
- In terms of equation, (9-10), calculate BDF.
- Sort column vector of image candidate features by descend, then select top n maximal value and their corresponding position.
- Select top n maximal value of every column vector as candidate features, obtain feature matrix of images.
- Calculate the distance from feature projection of test images and training feature matrix. In terms of the minimal distance, recognize images.

4. Experiment and Result Analysis

To illustrate the validity of algorithms for face recognition, the proposed method is compared with related methods in Yale and Yale B face database. Furthermore, we discussed its performance and efficiency.

4.1. Experiment Result

(1) Experiment on Yale database

To reduce computational cost and ensure recognition rate, the images in database are down-sampled into 64×64 pixels. Every image is divided into 8*8 blocks, which per block are 8×8 pixels. With DCT, every block has 64 coefficients. DC and top 3 AC components are used as candidate features. Hence, number of total candidate features is 64×4=256.

For testing, the cross validation method is used. 3-5 labeled images are selected as the training sample respectively; the others are selected as testing samples. The results are average of 20 runs.

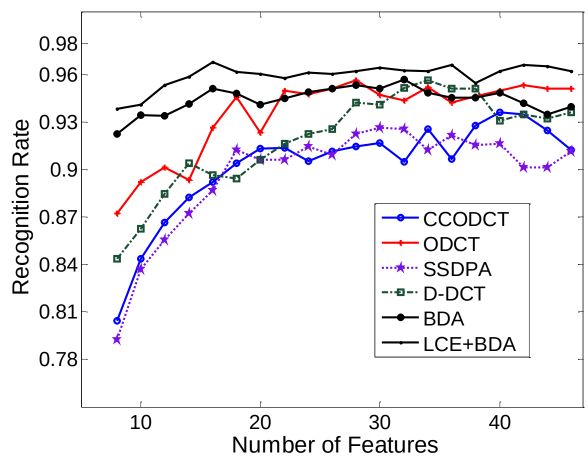

Figure 7. Recognition Rate of 6 Different Methods on Yale

Figure 7 illustrates recognition rate of 6 different methods on Yale face database. 5 images are selected as training samples, 6 different methods such as BDA, LCE+BDA, ODCT, SSDPA, D-DCT and CCODCT are used to recognize face images. It shows the recognition rate of LCE+BDA and BDA outperform the other 4 methods. Furthermore, LCD+BDA obtains higher recognition rate than BDA. It proves LCE reduce deviation of samples from same subject and greatly enhance training efficiency.

Table 2 shows recognition rate of 10 related methods on Yale database, which 3, 4 and 5 training samples are selected from per subject. Eigenface is unsupervised learning method. Fisherface, DLDA, BDA, LCE+BDA and D-DCT are supervised learning methods. SDA, CCODCT and SSDPA are semi-supervised learning methods. We can see LCE+BDA obtains best recognition rate under different training samples. D-DCT, ODCT, CCODCT, SSDPA also obtained satisfactory recognition rate.

Table 2. Recognition Rate of Different Methods on Yale (%)

Method	Sample Number		
	i=3	i=4	i=5
Eigenface[6]	69.24	71.26	75.32
Fisherface[7]	73.24	77.28	81.26
DLDA[8]	74.28	79.26	82.14
SDA[15]	76.25	81.24	84.63
D-DCT[1]	82.15	85.32	90.14
ODCT[13]	83.26	86.17	92.12
CCODCT[13]	82.15	87.26	91.37
SSDPA[14]	83.15	85.68	92.15
BDA	86.29	88.48	94.14
LCE + BDA	88.36	93.15	96.15

(2) Experiment on Yale database Yale B

Yale B includes 5760 face images from 10 subjects. For each subject, there are 576 images, which has 9 postures under 64 illumination conditions. To test the robustness on illumination variation of proposed methods, we use a subset of Yale B, which includes 640 frontal facial images from 10 subjects under same posture. Similar to Yale database, each image in the center area are down-sampled into 64×64 pixels. Down-sampled image is divided into 8*8 blocks, which are processed by feature extraction methods same as Yale. Hence, candidate features is 8*8*4=256. 15, 20, 30 labeled images are selected as the training sample respectively, and the others are used for testing. As fixed postures from every subject are selected, purpose of testing is to validate robustness on illumination of different methods.

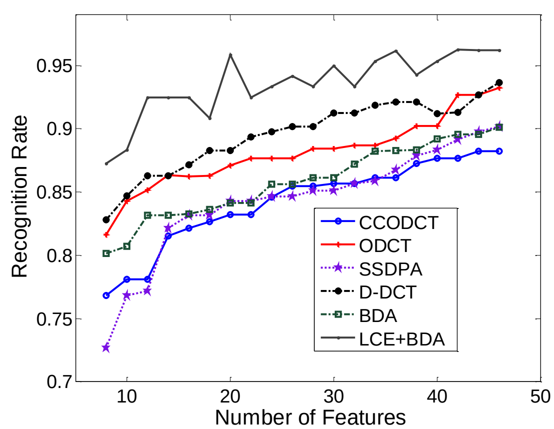

Figure 8. Recognition Rate of 3 DCT methods and combination on Yale B

Figure 8 illustrates 20 images from per subject are used to be training samples, recognition rate of 6 methods which are BDA, LCE+BDA, ODCT, SSDPA, D-DCT and CCODCT. LCE+BDA obviously out per forms other unused LCE methods. It shows LCE can compensate for uneven illumination and enhance texture of image to improve recognition rate. BDA also obtains better recognition rate than D-DCT、ODCT、CCODCT and SSDPA. It shows block DCT can extract important feature information.

Table 3 lists recognition rate of related methods on fixed posture under illumination variation condition from Yale B. 15, 20 and 30-labeled training samples are selected, and the other images are used for testing. Among them, LCE+BDA also obtain the best recognition rate.

Table 3. Recognition Rate of Different Methods on Yale B (%)

Method	Sample Number		
	i=15	i=20	i=30
Eigenface	68.47	71.28	76.28
Fisherface	71.26	76.24	81.25
DLDA	73.26	77.24	82.34
SDA	71.26	75.39	83.36
D-DCT	76.36	82.14	85.62
ODCT	79.12	87.12	91.16
CCODCT	77.38	86.16	88.69
SSDPA	76.67	87.12	89.26
BDA	82.15	89.15	93.62
LCE+ BDA	84.68	91.24	94.95

4.2 Evaluation on Performance

(1) Analysis on efficiency

Table 4 lists computational time of 10 methods. We can see the computational time of LCE+BDA is longer than BDA. Because its time is made up of two parts, which include preprocessing time and feature extraction time, while BDA do not include preprocessing time. The computational time of SSDPA and ODCT is shorter than BDA. Though candidate featues is fewer, the procedure of dividing block need some time. As eigen value computation need much time, computational time of Eigen face and Fisher face is bigger than other methods.

Table 4. Compute Time of Different Methods (ms /per Image)

Method	Database	
	Yale	YaleB
Eigenface	1.94	1.87
Fisherface	2.14	2.03
DLDA	1.78	1.68
SDA	1.61	1.54
D-DCT	1.48	1.39

ODCT	0.48	0.42
CCODCT	1.29	1.25
SSDPA	0.43	0.46
BDA	0.52	0.59
LCE+ BDA	0.86	0.93

(2) Evaluation on performance

With comparison of different methods on Yale and Yale B, recognition rate of LCE+BDA outperforms other listed methods. Because its processing step is easy and short computational time, the proposed methods get best performance.

5. Conclusion and Future Work

In the paper, LCE+BDA are proposed for face recognition. Its advantages are as follows: enhances local tiny texture by LCE, and obtains good robustness on illumination variation; As it doesn't need to compute eigen value, small sample problem can be solved; Without using complex functions, it has low computational cost. However, there are still many problems to be solved, such as how to capture labeled samples, recognizing for side facial image and partial occluded image etc. Next work is to apply this method to semi-supervised learning methods with few labeled samples.

Acknowledgements

This work was Supported by Scientific Research Fund of Heilongjiang Provincial Education Department(NO:11551086).

References

[1] X. Y. Jing and D. Zhang, "A face and palmprint recognition approach based on discriminant DCT feature extraction", IEEE Trans. on Syst. Man Cybern B, vol. 34, no. 6, **(2004),** pp. 2405-2415.
[2] M. J. Er, W. Chen and S. Wu, "High-speed face recognition based on discrete cosine transform and RBF neural networks", IEEE Trans. on Neural Networks, vol. 16, no. 3, **(2005),** pp. 679-691.
[3] F. Wang, J. Wang, C. Zhang and J. Kwok, "Face recognition using spectral features", Pattern Recognition, vol. 40, no.7, **(2007)**, pp. 2786-2797.
[4] N. Kwak, "Feature extraction for classification problems and its application to face recognition", Pattern Recognition, vol. 41, no. 5, **(2008)**, pp. 1718-1734.
[5] Z. Liu and C. Liu, "A hybrid color and frequency features method for face recognition", IEEE Trans. Image Process, vol. 17, no. 10, **(2008)**, pp. 1975-1980.
[6] M. Kirby and L. Sirovich, "Application of the Karhunen–Loeve procedure for the characterization of human faces", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 1, **(1990)**, pp. 103–108.
[7] M. Turk and A. Pentland, "Eigenfaces for recognition", International Journal of Cognitive Neuroscience, vol. 3, no. 1, **(1991)**, pp. 71-86.
[8] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data with application to face recognition", Pattern Recognition, vol. 34, no. 10, **(2001)**, pp. 2067-2070.
[9] J. Y. Kim, L. S. Kim and S. H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization", IEEE Trans. Circuits Syst. Video Technol., vol. 11, no. 4, **(2001)**, pp. 475-484.
[10] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition", International Conference on Audio and Video-Based Biometric Person Authentication, **(2003)**, pp. 10-18.
[11] W. C. Kao, M. C. Hsu and Y. Y. Yang, "Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition", Pattern Recognition, vol. 43, no. 5, **(2010)**, pp. 1736-1747.
[12] S. Dabbaghchian, M. P. Ghaemmaghami and A. Aghagolzadeh and A Aghagolzadeh, "Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology", Pattern Recognition, vol. 43, no. 4, **(2010)**, pp. 1431-1440.

[13] P. Cui and R. B. Zhang, "A semi-supervised coefficient select method for face recognition", Journal of harbin engineering university, vol. 33, no. 7, **(2012)**, pp. 855-861.
[14] P. Cui and R. B. Zhang, "Face feature extraction method based on part of labeled data", Journal of optoelectronics laser, vol. 23, no. 3, **(2012)**, pp. 554-560.
[15] D. Cai, X. F. He and J. W. Han, "Semi-supervised discriminant analysis", 2007 IEEE 11th International Conference on Computer Vision, **(2007)**, pp. 205-211.

Authors

Peng Cui, he was born in Heilongjiang, China, in 1971. He received the M. S. degree from Harbin University of Science and Technology in 2004, and received the Ph. D. degreee from Harbin University of Engineering in 2012. Now he is a professor. Currently his main research interests are image processing and pattern recognition.

