Acta Technica 62 (2017), No. 5B, 911–920

Face illumination processing based on combined features

Ming Liu[1], Tao Xie[2,3]

Abstract. Considering the significant impact of illumination variation on the face recognition, this paper presents a method for processing face illumination based on combined features. It combines two different methods of processing face illumination and further uses the principle component analysis (PCA) to reduce the dimensionality. The obtained results are then sorted by column and classified by the nearest neighbor classifier based on the cosine distance. The Yale-B face database and CMU-PIE face database are used to test the effectiveness of the combination of discrete cosine transform (DCT) and gradient faces and the effectiveness of the combination of wavelet transform and gradient faces. The results show that the effect of feature combination is better than that of the separate method.

Key words. discrete cosine transformation, gradient faces, wavelet transform, combined features, face recognition.

1. Introduction

With the vigorous development of science and technology, the face recognition has become one of the hot research focuses in the field of computer vision. The current face recognition techniques are facing many challenges, including the variations of illumination, posture, and expression, wherein the illumination variation poses a more severe hardship than others. The difference of the same face image under different illumination conditions is sometimes much greater than the difference of diversified face images under the same illumination condition [1]. In recent years, many scholars have studied the problem of illumination variation in face recognition. The studies can be classified into three categories, including the method of feature extraction based on the invariability of human face illumination (FEI), the method based on face modeling (FM), and the method of illumination preprocessing and

[1]Workshop 1 - Computer College, Chongqing College of Electronic Engineering, Chongqing, China; e-mail: 260866996@qq.com

[2]Workshop 2 - College of Computer and Information Science, Southwest University, Chongqing, China; e-mail: grabertao@gmail.com

[3]Corresponding author: Tao Xie; e-mail: grabertao@gmail.com

normalization (IPN). The main idea of the FEI method is to extract the insensitive features of illumination of face images. Examples of this method are the edge map approach, quotient image approach (QI), self-quotient image approach and local binary pattern approach (LBP) [2]. The FM method is to generate a face model which can recombine face images with different gestures and different illumination conditions. Examples of this method include the principal component analysis (PCA) [3], linear discriminant analysis (LDA) [4], illumination cone [5] and spherical harmonics [6]. The main limitation is that this kind of method requires a large number of face images under different lighting conditions as training samples, which is not practical in real applications. The core of the IPN method is to convert the image into a standard form before the face recognition, and further use it for face discrimination. Examples of this method are the histogram equalization [7], gamma correction and logarithmic transformation [8] etc. But these algorithms cannot effectively overcome the problem of face variation that caused by the local illumination.

This paper presents a method for processing face illumination based on combined features. The framework of the algorithm is shown in Fig. 1. The method takes respective advantage of different combinations of illumination processing, uses proper light compensations on the images and further runs the PCA-based dimensionality reduction. In this paper, the authors respectively use the combination of DCT and gradient faces (combination 1) and the combination of wavelet transform and gradient faces (combination 2). The results are grouped by column and processed with the nearest neighbor classifier based on the cosine distance. The experiments are conducted on the Yale-B and CMU-PIE databases.

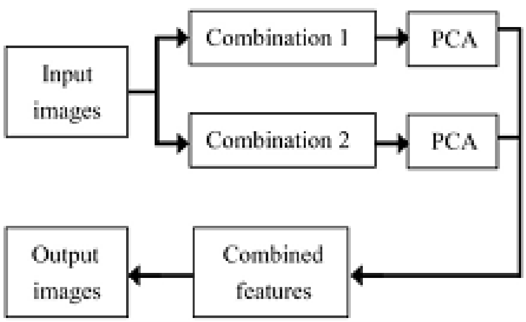

Fig. 1. Framework of the proposed algorithm

2. The proposed algorithm

2.1. Wavelet transform of images

Given a multiresolution analysis (MRA) $\{V_{\mathrm{j}}\}_{\mathrm{j}\in\mathrm{Z}}$ on $L^2(R)$ and a corresponding scaling function $\phi(x)$, the authors define a two-dimensional scale space $V_j^2 = V_j \otimes V_j$ of scale j, where $\otimes$ denotes the space multiplying operation. Because the standard orthogonal basis of V_j can be represented as

$$\phi_{\mathrm{j,n}}(x) = 2^{-\mathrm{j}/2}\phi\left(2^{-\mathrm{j}}x - n\right) \tag{1}$$

The authors can obtain the standard orthogonal basis of V_{j}^2, represented as $\{\phi_{\mathrm{j,n}}(x)\,\phi_{\mathrm{j,m}}(x)\}_{\mathrm{n,m}\in\mathrm{Z}}$.

Let $W_{\mathrm{j+1}}^2$ denote the orthogonal complement space of $V_{\mathrm{j+1}}^2$ and V_{j}^2, and let $\psi(x)$ denote the one-dimension wavelet of $\phi(x)$. The authors can obtain three wavelet functions of $L^2\left(R^2\right)$ which make $\left\{\Psi^{\mathrm{i}}\left(2^{\mathrm{j}}x - k, 2^{\mathrm{j}}y - l\right)\right\}$, $j, k, l \in Z, i = 1, 2, 3$ to be the orthonormal basis of $L^2\left(R^2\right)$. The three wavelet functions are

$$\begin{cases} \Psi^1 = \phi \otimes \psi = \phi(x)\,\psi(y) \\ \Psi^2 = \psi \otimes \phi = \psi(x)\,\phi(y) \\ \Psi^3 = \psi \otimes \psi = \psi(x)\,\psi(y) \end{cases} \tag{2}$$

In two dimensional MRA, for the squared-integrable two-dimensional image signals $f(m,n) \in L^2\left(R^2\right)$, let $C_{0,\mathrm{mn}} = f(m,n)$, the wavelet decomposition of the image can be recursively represented as:

$$\begin{cases} C_{j,mn} = \sum_{k,l} C_{j-1,kl}h_{k-2m}h_{l-2n} \\ D_{j,mn}^1 = \sum_{k,l} C_{j-1,kl}h_{k-2m}g_{l-2n} \\ D_{j,mn}^2 = \sum_{k,l} C_{j-1,kl}g_{k-2m}h_{l-2n} \\ D_{j,mn}^3 = \sum_{k,l} C_{j-1,kl}g_{k-2m}g_{l-2n} \end{cases} \tag{3}$$

where $h(n) =< \phi, \phi_{-1,n} >$ and $g(n) =< \psi, \phi_{-1,n} >$. Also, the authors have

$$g(n) = (-1)^n\, h(1-n) \tag{4}$$

After the level-1 wavelet decomposition, the level-2 wavelet decomposition generates four sub images, respectively the low-frequency component $LF1$, the horizontal detail component $HD1$, the vertical detail component $VD1$ and the diagonal detail component $DD1$. The sizes of the four components are half of the raw images. If the images are required to go through the level-2 wavelet decomposition, the authors further decompose the $LF1$ component into the low-frequency component $LF2$ and three detail components, i.e. the horizontal detail component $HD2$, the vertical detail component $VD2$ and the diagonal detail component $DD2$. It is worth noting that the sizes of the components are half of the size of $LF1$. If the images are required to go through the next level decomposition, the authors repeat the similar ways to decompose $LF2$, and so forth. Finally, the sub-image $LF\mathrm{n}$ obtained by level-n wavelet decomposition of the image is the low frequency approximation of the raw image.

2.2. Algorithm for gradient faces

The computation of gradient faces is below.

The grayscale images of faces can be treated as the product of reflection components $I(x,y)$ and illumination components $L(x,y)$:

$$I(x,y) = R(x,y) \times L(x,y) \tag{5}$$

where $R(x,y)$ corresponds to the fast-changing part of the image and $L(x,y)$ corresponds to the slow-changing part of the image.

For grayscale images $I(x,y)$??the neighborhood point of pixel point (x,y) along x direction is $(x+\Delta x, y)$. Then, the authors have [9]:

$$I(x+\Delta x, y) = R(x+\Delta x, y) L(x+\Delta x, y) \tag{6}$$

Subtracting the Equation (5) from the Equation (6), the authors obtain:

$$I(x+\Delta x,y) - I(x,y) = R(x+\Delta x,y)L(x+\Delta x,y) - R(x,y)L(x,y) \tag{7}$$

Because L is smooth and almost constant, then:

$$I(x+\Delta x, y) - I(x,y) \approx [R(x+\Delta x) - R(x,y)]\, L(x,y) \tag{8}$$

When Δx tends to be 0, the Equation (8) is equivalent to:

$$\frac{\partial I(x,y)}{\partial x} \approx L(x,y)\frac{\partial R(x,y)}{\partial x} \tag{9}$$

Similarly, in the y direction, the authors have:

$$\frac{\partial I(x,y)}{\partial y} \approx L(x,y)\frac{\partial R(x,y)}{\partial y} \tag{10}$$

The Equation (10) divided by the Equation (9), the authors obtain??

$$\frac{\partial I(x,y)}{\partial y}\bigg/\frac{\partial I(x,y)}{\partial x} \approx \frac{\partial R(x,y)}{\partial y}\bigg/\frac{\partial R(x,y)}{\partial x} \tag{11}$$

The gradient face is defined as:

$$G = \arctan(Iy/Ix) \tag{12}$$

where $I_y = \partial I(x,y)/\partial y$ and $I_x = \partial I(x,y)/\partial x$ are the respective gradients along the x- and y- directions, $G \in [0, 2\pi)$. To get the gradient faces in practical applications, the authors need calculate the gradient of the face images in the x- and y- directions. For this purpose, the authors use the Gaussian kernel function to smooth the face image, convolve the faces and the Gaussian kernel, and then compute the gradient along the x- and y- directions separately. The Gaussian kernel function is computed by:

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{13}$$

where σ is the standard deviation of the Gaussian kernel function.

2.3. Method of feature combination

The feature combination takes the respective advantage of different methods of illumination processing and uses the combined features to reduces the influences of illumination on the face recognition. The authors respectively use the combination of discrete cosine transform (DCT) and gradient faces and the combination of wavelet transform and gradient faces. The realization of this process is as follows:

Input: raw face image I;

Output: processed face feature $I^{'}$.

1.Perform logarithmic transformation on I;

2. Discard the low-frequency coefficients of I with DCT and obtain I_1;

The DCT of a discrete function $f(x), x = 0, 1, ..., N-1$ is defined as:

$$F(k) = \frac{2ck}{N}\sum_{x=0}^{N-1} f(x)\cos\left[\frac{(2x+1)k\pi}{2N}\right] \tag{14}$$

where $k = 0, 1, ..., N-1$.

For an $M \times N$ image, where each image corresponds to a two-dimensional matrix, the DCT coefficients are calculated as:

$$F(u,v) = \frac{1}{\sqrt{MN}}\alpha(u)\alpha(v) \times \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{(2x+1)u\pi}{2M}\right]\cos\left[\frac{(2y+1)v\pi}{2N}\right] \tag{15}$$

where $u = 0, 1, ..., N-1$ and $v = 0, 1, ..., N-1$.

The $\alpha(\omega)$ is defined as:

$$\alpha(\omega) = \begin{cases} 1/\sqrt{2}, & for\ \ \omega = 0 \\ 1, & for\ \ \omega = 1, 2, ..., N-1 \end{cases} \tag{16}$$

3. Use the gradient face algorithm presented in section 2.2 to enhance the high-frequency part and obtain I_2;

4. Run PCA on I_1 and I_2 to respectively get the projections $I_1^{'}$ and $I_2^{'}$;

The Singular Value Decomposition (SVD) is used to extract the eigenvalues and eigenvectors. Let r be the number of principal components which takes the vast majority of energy of the original matrix. The authors compute the cumulative energy ratio no less than a cutoff of 95% to determine r .

$$r = \arg\max_r\left(\frac{\sum_1^r \lambda_i}{\sum_1^n \lambda_i} \times 100\%\right) \geq 95\% \tag{17}$$

where λ is the eigenvalue.

5. Use $I' = \begin{bmatrix} I_1^{'} \\ I_2^{'} \end{bmatrix}$ combination to obtain light-insensitive facial features.

3. Experiment

3.1. Description of face databases

The authors conduct a series of experiments on the Yale-B and CUM-PIE databases. The Yale-B database contains nine postures of ten individual persons, and each posture contains 64 illumination conditions. Since the authors deal with the illumination process problem, only the face images of positive posture are selected in the experiment. The image size is 128×128, and all face images are divided into five subsets according to different angles of incident light. The subset 1 ($\theta < 12°$) has 70 samples, the subset 2 ($13° < \theta < 25°$) has 120 samples, the subset 3 ($26° < \theta < 50°$) has 120 samples, the subset 4 ($51° < \theta < 77°$) has 140 samples, and the subset 5 ($\theta > 77°$) has 190 samples. The subset 1 is used for training, and the others are used for testing.

The CUM-PIE database includes 41,368 face images of 68 persons. Similar to Yale-B, the authors select the light group (C27) of positive face as a training gallery, which contains 1,428 faces in total. All images are cropped to 64 x 64 sizes. In the experiment, the authors select a face image from 21 illuminations of each sample and use it as the training data with the remaining being the test data.

3.2. Parameter setting

The standard deviation σ in the Equation (**??**) should be valued in a proper interval to get the best recognition rate. Also, the authors select the different wavelet types and decomposition series to study how they influence the results.

3.3. Experiment results

The table 1 shows the number of discarded low frequencies of DCT and the data range of standard deviation σ of Gaussian function in Equation (**??**) when the recognition accuracy reaches 100%. The table 2 shows the wavelet type, decomposition series and data range of standard deviation σ of Gaussian function. The results of comparison on Yale-B are presented in table 3.

Table 1. The DCT is combined with the gradient faces (Yale-B)

The number of discarded frequencies of DCT	16	17	40
Range of σ	0.7-0.8	0.1-1.0, 1.2	1.0-1.1

Table 2. The wavelet transform is combined with the gradient faces (Yale-B)

Wavelet type (series)	Range of σ	Wavelet type (series)	Range of σ
db1(3)	1.3	db10(3)	1.0-1.1,1.3
db6(3)	0.1-0.8	bior5.5(3)	0.1-0.8
db10(3)	0.1-0.8	bior5.5(3)	1.0-1.6

From the table 1, the impact of illumination on the face recognition can be effectively reduced if the authors select proper DCT coefficients and the range of σ when the DCT is combined with the gradient faces, which makes the recognition accuracy on Yale-B database get 100%. Also, the authors can learn from the table 2 that the level-3 decomposition using different ranges of σ and wavelet types obtain the best performance when the wavelet transform is combined with the gradient faces.

The table 3 presents the comparisons of different methods, including the method proposed in this paper, DCT, gradient faces [10], LBP [11] and QI [12] under the same experimental conditions.

To further validate the effectiveness of the proposed method, the authors also conduct experiments on the CUM-PIE database. The results are shown in table 4. The table 4 compares our method with the separate DCT, wavelet transform, gradient faces, and multiscale retinex (MSR) . For convenience, the combination of DCT and gradient faces is called the Combination 1, and the combination of wavelet transform and gradient faces is called the Combination 2.

Table 3. Accuracy of recognition of different approaches on Yale-B

Approaches	Accuracy of recognition (%)				
	Subset 2	Subset 3	Subset 4	Subset 5	average
DCT	0	99.17	97.86	97.37	98.42
Gradient faces	0	0	98.57	98.95	99.3
LBP	0	97.6	65.2	44.4	76.8
QI	99.3	61.9	34.1	23.3	54.65
Combination 1	98.43	99.39	98.1	97.98	99.23
Combination 2	97.95	99.54	99.19	98.10	99.84

Table 4. Accuracy of recognition of different approaches on CUM-PIE

	Approaches					
	DCT	Gradient faces	Wavelet transform	MSR	Combination 1	Combination 2
Accuracy (%)	94.41	95.07	93.53	84.50	96.91	96.69

The experiment conducted on CUM-PIE database shows that the recognition is the best when the number of discarded DCT coefficients is 19 while using the DCT independently. It also shows that the best recognition is achieved when $\sigma = 0.55$ while using the gradient faces independently. Furthermore, when the wavelet type is db10, the effect of level-2 decomposition is very significant. The combination of DCT and gradient faces gets the highest accuracy when $\sigma = 0.7$ and the number of discarded DCT coefficients is 10.

The table 4 is showing that the combined method proposed in this paper is superior to any separate algorithm. Furthermore, the proposed method can effectively reduce the influence of different angles on the face image compared to other existing algorithms.

4. Conclusion

This paper presents a method of face illumination based on feature combination. Firstly, the face images are processed by two different illumination algorithms, and then the processed images are extracted from the face features using PCA. The results are combined by column, and the nearest neighbor classifier based on cosine distance is used to discriminate the class label. The experimental results show that in the positive Yale-B face database, the combination of DCT and gradient faces and the combination of wavelet transform and gradient faces can almost get 100% recognition accuracy by appropriately selecting discarded DCT coefficients, Gaussian kernel function, wavelet type and the decomposition series. Also, the recognition performance on the CUM-PIE face database is better than that of using a separate method and other existing methods. The recognition rate of our method can reach more than 95%. Therefore, the face illumination processing algorithm proposed in this paper has a good robustness to the illumination, which can effectively reduce the influence of different angles on the face image.

References

[1] C. Kaymak,et al: *Illumination Invariant Face Recognition Using Principal Component Analysis–An Overview.* Machine Vision and Mechatronics in Practice, Springer Berlin Heidelberg (2015), 269-285.

[2] H. Tang, et al: *3D face recognition using local binary patterns.* Signal Processing *93* (2013), No. 8, 2190-2198

[3] M. Imran, et al: *Face Recognition using Eigenfaces.* International Journal of Computer Applications *118* (2015), No. 5.

[4] P. N. Belhumeur, et al: *Eigenfaces vs. fisherfaces: Recognition using class specific linear projection.* IEEE Transactions on pattern analysis and machine intelligence *19* (1997), No. 7, 711-720

[5] A. S. Georghiades, et al: *From few to many: Illumination cone models for face recognition under variable lighting and pose.* IEEE transactions on pattern analysis and machine intelligence *23* (2001), No. 6, 643-660.

[6] Z. Khalid, et al: *An optimal-dimensionality sampling scheme on the sphere with fast spherical harmonic transforms.*IEEE Transactions on Signal Processing *5* (2014),No. 17, 4597-4610.

[7] J. R. Tang, N. A. M. Isa: *Bi-histogram equalization using modified histogram bins.* Applied Soft Computing *55* (2017), 31-43.

[8] Y. Xu, et al: *Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification.* Pattern Recognition *54* (2016), 68-82.

[9] Z. Li, et al: *Score level fusion method based on multiple oblique gradient operators for face recognition.* Multimedia Tools and Applications *75* (2016), No. 2, 819-837.

[10] X. Chen, J. Zhang: *Illumination robust single sample face recognition using multi-directional orthogonal gradient phase faces.* Neurocomputing *74*, (211), No. 14, 2291-2298.

[11] B. Zhang, et al: *Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor.* IEEE transactions on image processing *19* (2010), No. 2, 533-544.

[12] A. Shashua, T. Riklin-Raviv: *The quotient image: Class-based re-rendering and recognition with varying illuminations.* IEEE Transactions on Pattern Analysis and Machine Intelligence *23* (2001), No. 2, 129-139.

Received November 16, 2016