A Low-dimensional Illumination Space Representation of Human Faces for Arbitrary Lighting Conditions

HU Yuan-Kui[1] WANG Zeng-Fu[1]

Abstract The proposed method for low-dimensional illumination space representation (LDISR) of human faces can not only synthesize a virtual face image when given lighting conditions but also estimate lighting conditions when given a face image. The LDISR is based on the observation that 9 basis point light sources can represent almost arbitrary lighting conditions for face recognition application and different human faces have a similar LDISR. The principal component analysis (PCA) and the nearest neighbor clustering method are adopted to obtain the 9 basis point light sources. The 9 basis images under the 9 basis point light sources are then used to construct an LDISR which can represent almost all face images under arbitrary lighting conditions. Illumination ratio image (IRI) is employed to generate virtual face images under different illuminations. The LDISR obtained from face images of one person can be used for other people. Experimental results on image reconstruction and face recognition indicate the efficiency of LDISR.

Key words LDISR, basis image, illumination ratio image, face recognition

1 Introduction

Illumination variation is one of the most important factors which reduce significantly the performance of face recognition system. It has been proved that the variations between images of the same face due to illumination are almost always larger than image variations due to change in face identity[1]. So eliminating the effects due to illumination variations relates directly to the performance and practicality of face recognition system.

To handle face image variations due to changes in lighting conditions, many methods have been proposed thus far. Generally, the approaches to cope with variation in appearance due to illumination fall into three kinds[2]: invariant features, such as edge maps, images filtered with 2D Gabor-like functions, derivatives of the gray-level image, images with Log transformations and the recently reported quotient image[3] and self-quotient image[4]; variation-modeling, such as subspace methods[5~7], illumination cone[8~10]; and canonical forms, such as methods in [11, 12].

This paper investigates the subspace methods for illumination representation. Hallinan *et al.*[5,6] proposed an eigen subspace method for face representation. This method firstly collected frontal face images of the same person under different illuminations as training set, and then used principal component analysis (PCA) method to get the eigenvalues and eigenvectors of the training set. They concluded that 5 ± 2 eigenvectors would suffice to model frontal face images under arbitrary illuminations. The experimental results indicated that this method can reconstruct frontal face images with variant lightings using a few eigenvectors. Different from Hallinan, Shashua[7] proposed that under the assumption of Lambertian surface, three basis images shot under three linearly independent light sources could reconstruct frontal face images under arbitrary lightings. This method was proposed to discount the lighting effects but not to explain lighting conditions. Belhumeur *et al.*[8,9] proved that face images with the same pose under different illumination conditions form a convex cone, called illumination cone, and the cone can be represented in a 9 dimensional space[10]. This method performs well but it needs no less than seven face images for each person to estimate the 3D face shape and the irradiance map. Basri & Jacobs [13] and Ramamoorthi[14,15] independently applied the spherical harmonic representation and explained the low dimensionality of differently illuminated face images. They theoretically proved that the images of a convex Lambertian object obtained under a wide variety of lighting conditions can be approximated accurately with a 9D linear subspace, explaining prior empirical results[5~7]. However, both of them assumed that the 3D surface normal and albedo (or unit albedo) were known. This assumption limits the application of this algorithm.

The above research results theoretically and empirically indicate that frontal face images obtained under a wide variety of lighting conditions can be approximated accurately with a low-dimensional linear subspace. However, all the above subspace methods construct a subspace from training images for each human face, which is not only corresponding to the illumination conditions but also to the face identity. The subspaces, in which the intrinsic information (shape and albedo) and the extrinsic information (lightings) are mixed, are not corresponding to the lighting conditions distinctly. Otherwise, a large training image set would be needed in the learning stage and 3d face model might be needed.

In this paper, a low-dimensional illumination space representation (LDISR) of human faces for arbitrary lighting conditions is proposed, which can handle the problems that can not be solved well in the existing methods to a certain extent. The key idea underlying our model is that any lighting condition can be represented by 9 basis point light sources. The 9 basis images under the 9 basis point light sources construct an LDISR, which separates the intrinsic and the extrinsic information and can both estimate lighting conditions when given a face image and synthesize a virtual face image when given lighting condition combining with the illumination ratio image (IRI) method. The method in [10] and the proposed method in this paper have some similarities, but they have some essential differences also. The former needs to build one subspace for each person, and the latter only needs to build one subspace for one selected person. Furthermore, the 9D illumination space built in the former case is not corresponding to the lighting conditions distinctly, and in our case once the corresponding illumination space is built, it can be used to generate virtual frontal face images of anybody under arbitrary illuminations by using the warping technology and IRI method developed. These virtual images are then used for the purpose of both training and recognition. The experiments on

Received January 11, 2006; in revised form March 28, 2006
Supported by Open Foundation of National Laboratory of Pattern Recognition, P. R. China.
1. Department of Automation, University of Science and Technology of China, Hefei 230027, P. R. China
DOI: 10.1360/aas-007-0009

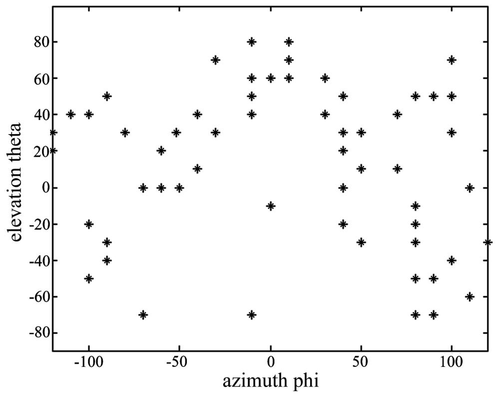

Fig. 1 The positions corresponding to the dominant point light sources

Yale Face Database B indicate that the proposed method can improve the performance of face recognition efficiently.

2 Constructing the LDISR

Since any given set of lighting conditions can be exactly expressed as a sum of point light sources, a surface patch′s radiance illuminated by two light sources is the sum of the corresponding radiances when the two light sources are applied separately. More detail was discussed in [5]. In this section, PCA and clustering based method are adopted to find the basis point light sources, which are able to represent arbitrary lighting conditions.

The needed 3D face model was obtained using a 3D imaging machine *3DMetrics* TM. Then the 3D face model obtained was used to generate the training images. Move a floodlight by increments of 10 degrees to each position (θ_i, φ_j) to generate image $\boldsymbol{p}(\theta_i, \varphi_j)$, where θ is the elevation and φ is the azimuth. Typically $\varphi \in [-120^\circ, 120^\circ]$ and $\theta \in [-90^\circ, 90^\circ]$. Totally, 427 images were generated, denoted as $\{\boldsymbol{p}_k, k = 1, \cdots, 427\}$.

We use PCA to find the dominant components for the finite set of images. Since the PCA is used on the images of the same human face with different lighting conditions, the dominant eigenvectors do not reflect the facial shape but the lighting conditions. So the above eigenvectors can be used to represent lighting conditions. In this paper, the lighting subspace is constructed not using the eigenvectors directly but the light sources corresponding to the eigenvectors.

According to the ratio of the corresponding eigenvalue to the sum of all the eigenvalues, the first 60 eigenvalues containing the 99.9% energy were selected. And the 60 corresponding eigenvectors were selected as the principal components. Denote the first 60 eigenvectors as $\{\boldsymbol{u}_i, i = 1, \cdots, 60\}$. For the ith eigenvector $\boldsymbol{u}_i$, the corresponding training image is $\boldsymbol{p}_j$, where $\boldsymbol{u}_i$ and $\boldsymbol{p}_j$ satisfy

$$\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{p}_j = \max_{k \in \{1, \cdots, 427\}} \{\boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{p}_k\} \tag{1}$$

The positions of the 60 dominant point light sources are shown in Fig.1.

By investigating the positions of the dominant point light sources, it can be found that the dominant point light sources are distributed by certain rules. They are distributed almost symmetrically and cluster together in regions such as the frontal, the side, the below, and the above of head. The nearest neighbor clustering method is adopted here to get the basis light positions. Considering the effects of point light sources in different elevation and azimuth, some rules are employed for clustering:

1. When the elevation is below -60° or above 60°, clustering is done based on the differences of values in elevation.
2. When the elevation is in range $[-60^\circ, 60^\circ]$, clustering is done based on the Euclidian distances in space.

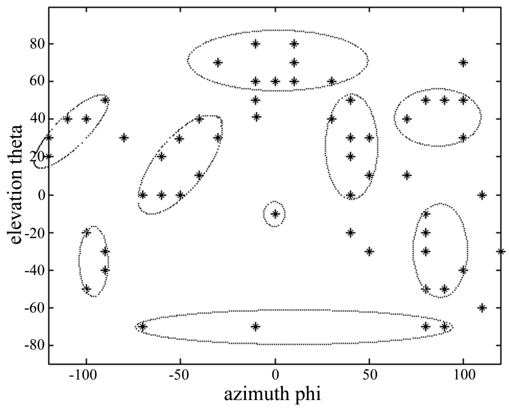

Fig. 2 The clustering result of the first 60 eigenvectors.

By adopting the nearest nerghbor clustering method, the 60 dominant light sources can be classified into 9 classes. The clustering result is shown in Fig.2. When the geometric center of each class is regarded as the basis position, the 9 basis light positions are shown in Table 1.

From the above procedure, it is known that point light sources in the 9 basis positions are dominant and principal components in the lighting space, and they can express arbitrary lighting conditions. The 9 basis images obtained under the 9 basis point light sources respectively construct a low-dimensional illumination space representation (LDISR) of human face, which can express frontal face images under arbitrary illuminations. Because different human faces have similar 3D shapes [3,16], the LDISR of different faces is also similar. As an approximation, it can be assumed that different persons have the same LDISR, which has been discussed in [17].

Denote the 9 basis images obtained under 9 basis lights are $I_i, i = 1, \cdots, 9$, the LDISR of human face can be denoted as $A = [I_1, I_2, \cdots, I_9]$. The face image under lighting s_x can be expressed as

$$I_x = A\boldsymbol{\lambda} \tag{2}$$

where $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \cdots, \lambda_9]^T$,$0 \le \lambda_i \le 1$ is the lighting parameters of image I_x and can be calculated by minimizing the energy function $E(\boldsymbol{\lambda})$:

$$E(\boldsymbol{\lambda}) = \|A\boldsymbol{\lambda} - I_x\|^2 \tag{3}$$

So we can get

$$\boldsymbol{\lambda} = A^+ I_x \tag{4}$$

where

$$A^+ = (A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}} \tag{5}$$

Table 1 Positions of the 9 basis light sources

light	1	2	3	4	5	6	7	8	9
Elevation θ (degree)	0	17.5	25.7	36	44	68.6	-33.3	-35	-70
Elevation φ (degree)	0	-47.5	44.4	-108	88	-3	85	-95	22.5

(a) Input image (b) ASM alignment (c) Warped mean shape

(d) The virtual images generated under different lightings

Fig. 3 Generating virtual images using the 9D illumination space and the IRI

Given an image of human face for learning images, the lighting parameters $\boldsymbol{\lambda}$ can be calculated by (4), and the virtual face images can be generated by (2) by using the lighting condition $\boldsymbol{\lambda}$. In order to use the LDISR learned from one human face to generate virtual images of other human faces, the illumination ratio image (IRI) based method is adopted in next section.

3 Generating virtual images using illumination ratio-image (IRI)

Denote the light sources as s_i, $i = 0, 1, 2, \cdots$, respectively, where s_0 is the normal light source, and I_{ji} the image under light source s_i for the person with index j. The IRI is based on the assumption that a face is a convex surface with a Lambertian function. A face image can be described as

$$I(u,v) = \rho(u,v)\boldsymbol{n}(u,v)\cdot\boldsymbol{l} \tag{6}$$

where, $\rho(u,v)$ is the albedo of point (u,v), $\boldsymbol{n}(u,v)$ is the surface normal at (u,v), and $\boldsymbol{l}$ is the direction of light.

Different from the quotient image[3], illumination ratio image is defined as follows[11,18,19,20].

$$R_i(u,v) = \frac{I_{ij}(u,v)}{I_{0j}(u,v)} \tag{7}$$

From (6) and (7), we have

$$R_i(u,v) = \frac{\rho_j(u,v)\boldsymbol{n}^{\mathrm{T}}(u,v)\cdot\boldsymbol{s}_i}{\rho_j(u,v)\boldsymbol{n}^{\mathrm{T}}(u,v)\cdot\boldsymbol{s}_0} = \frac{\boldsymbol{n}^{\mathrm{T}}(u,v)\cdot\boldsymbol{s}_i}{\boldsymbol{n}^{\mathrm{T}}(u,v)\cdot\boldsymbol{s}_0} \tag{8}$$

Equation (8) shows that the IRI can be determined only by the surface normal of a face and the light sources, which is independent of specific albedo. Since different human faces have the similar surface normal[3,16], the IRIs of different people under the same lighting condition can be considered to be the same. In order to eliminate the effect due to shapes of different faces, the following procedure should be done. Firstly, all faces can be warped to the same shape, and then the IRI is computed. In this paper, an ASM based method is used to perform the face alignment and all faces will then be warped to the predefined mean shape. After the procedure, all faces will have a quite similar 3D shape. That is to say, with the same illumination, IRI is the same for different people. The corresponding face image under arbitrary lighting condition can be generated from the IRI. Finally the face image is warped back to its original shape.

From (7), we have

$$I_{ij}(u,v) = I_{0j}(u,v)R_i(u,v) \tag{9}$$

Equation (9) means that, given the IRI under s_i and the face image under the normal lighting, we can relight the face under $\boldsymbol{s}_i$.

The face relighting problem can be defined as follows. Given one image, I_{a0} , of people A under the normal lighting $\boldsymbol{s}_0$, and one image, I_{bx} , of another people B under some specific lighting $\boldsymbol{s}_x$, how to generate the image, I_{ax}, of people A under lighting $\boldsymbol{S}_x$. Unlike [11, 18], the IRI under each lighting is unknown in this paper .

Given image I_{bx}, the IRI under lighting $\boldsymbol{s}_x$ can be calculated using the LDISR described in Section 2. Assume the LDISR, A , is learned from images of people M. The lighting parameter, $\boldsymbol{\lambda}_x$, of image I_{bx} is solved by the least-square method

$$A^{\mathrm{T}}A\boldsymbol{\lambda}_x = A^{\mathrm{T}}I_{bx} \tag{10}$$

$A\boldsymbol{\lambda}_x$ is the image of people M under lighting s_x, denoted as I_{mx}. The IRI under lighting $\boldsymbol{s}_x$ can be calculated by

$$R_x(u,v) = I_{xm}(u,v)/I_{0m}(u,v) \tag{11}$$

where I_{0m} is the image of people M under normal lighting. After the IRI under lighting $\boldsymbol{s}_x$ is calculated, the face image of people A can be relit under lighting $\boldsymbol{s}_x$ by $I_{xa}(u,v) = I_{0a}(u,v)R_x(u,v)$.

In general, given face image I_{0y} of arbitrary face Y under lighting $\boldsymbol{s}_0$, face image of Y under arbitrary lighting can be generated by the following procedure:

1. Detect face region I_{0y} and align it using ASM;

2. Warp I_{0y} to the mean shape T_0;

3. Relight T_0 using the IRI under lighting $\boldsymbol{s}_k$: $T_k(u,v) = T_0(u,v)R_k(u,v)$;

4. Reverse-warp the texture T_k to its original shape to get the relit image I_{ky}

Fig. 3 shows some relighting results on Yale Face Database B. In the experiments, the LDISR was constructed by the nine basis images of people DZF(not included in Yale Face Database B). For each image under

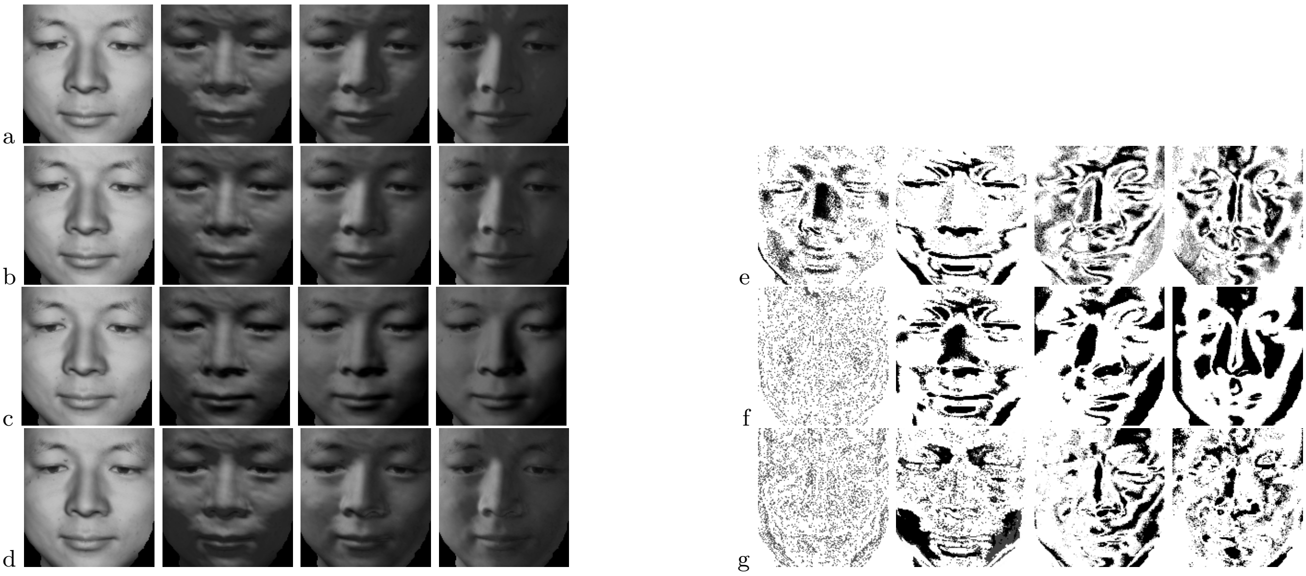

Fig. 4 Results of image reconstruction. a) Original images. b) Images reconstructed by 5-eigenimages. c) Images reconstructed by 3-basis images. d) Images reconstructed by the LDISR. e) The differences corresponding to the images in b). f) The differences corresponding to the images in c). g) The differences corresponding to the images in d).

normal lighting in Yale Face Database B, the virtual images under other 63 lightings were generated.

It should be highlighted that in the original IRI method[11,18], to calculate the IRI, the image under normal lighting and the image under specific lighting must be of the same people. The LDISR based method proposed in this paper breaks this limitation and the face image used in the algorithm can be of different people. In addition, when no face image under normal lighting is available, the virtual image can be generated by using the given λ_x from (2). And the IRI will then be calculated according to the virtual image.

4 Experimental results

4.1 2D image reconstruction

The experiment was based on the 427 frontal face images under different lightings described in Section 2. In this experiment, three image reconstruction methods were implemented: 5-eigenimages representation method proposed by Hallinan[5], a linear combination of 3-basis images proposed by Shashua[7], and the LDISR based method. The face images under different lightings were reconstructed and the performances were evaluated by the differences between the original and the reconstructed images.

According to [5], PCA was adopted to train the 427 images and the eigenvectors corresponding to the first 5 eigenvalues were selected to construct face illumination subspace I. According to [7], the selected 3 basis images under three point light sources respectively were used to construct face illumination subspace II. The LDISR constructed by the nine basis images was the face illumination subspace III. The total 427 face images were reconstructed by the three face illumination subspace, respectively.

Some original images are shown in Fig. 4 a), and the images reconstructed using face illumination I, II, III are shown in Fig. 4 b), c), and d), respectively. The corresponding differences are shown in Fig. 4 e), f), and g), respectively.

It can be concluded from Fig.4 that the performances of the 5-eigenimages representation method and the LDISR are comparative, and they are both better than that of the 3-basis images representation method. When the variation due to lighting condition is large (Fig. 4 c), columns 2, 3, and 4), the differences between the original and the reconstructed images are very large (Fig. 4 f), columns 2, 3, and 4), especially when there are shadows in face images.

To evaluate more rigorously, the fit function defined in [5] was adopted. The quality of the reconstruction can be measured by the goodness of the fit function:

$$\varepsilon = 1 - \frac{\| I_{rec} - I_{in} \|^2}{\| I_{in} \|^2} \tag{12}$$

where I_{rec} is the reconstructed image, and I_{in} is the original image. The values of the fit function corresponding to all the 427 reconstructions by three methods are shown in Fig. 5.

From Fig. 5, it can be seen that the fitness of images reconstructed by the 5-eigenimages representation method and the LDISR to the original image is very good, while the 3-basis images representation method is not so good. When the variation in lighting is larger (corresponding to the abscissas are 50 and 280 in Fig. 5) the performance of the LDISR is better than that of the 5-eigenimages representation method.

Besides, the 5-eigenimages and the 3-basis images representation methods need multiple images of each person, and train one model for each person. However, the LDISR trains one model using 9 basis images of one person, and can be used for other person by warping technique.

4.2 Face recognition with variant lightings based on virtual images

In this experiment, the LDISR and the IRI method were combined to generate virtual face images, which were used for face recognition with variant lightings. The experiments were based on the Yale Face Database B[10]. 64 frontal face images of each person under 64 different lightings were selected, and there were 640 images of 10 persons. The

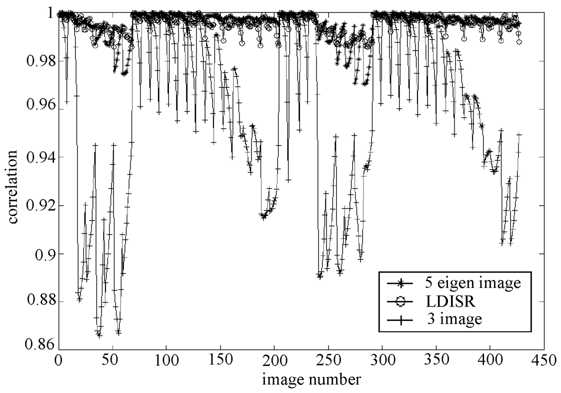

Fig. 5 The values of fit function corresponding to the reconstruction by three methods.

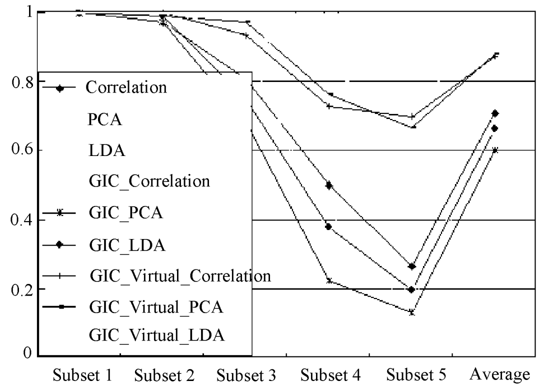

Fig.6 The results of Face recognition on Yale face Database B

images have been divided into five subsets according to the angles the light source direction makes with the camera axis[10]: Subset 1(up to 12°), subset 2 (up to 25°), subset 3(up to 50°), subset 4(up to 77°), and subset 5(others).

Correlation, PCA, and LDA methods were adopted for face recognition. For correlation method, the image under normal lighting of each person was the template image and the rest 63 images of each person were test images. For PCA and LDA methods, three images of each person (of which the angles the light source direction makes with the camera axis are the smallest) were training images, and the rest were test images.

The LDISR was constructed by the nine basis images of people DZF (not included in Yale Face Database B). For each frontal face image in Yale Face Database B, the virtual images corresponding to the other 63 lightings were generated using the LDISR and IRI. In order to decrease the effect of illumination, we used gamma intensity correction (GIC). Here $\gamma = 4$.

The three recognition methods were performed on the original images, images with GIC and virtual images with GIC. The results are shown in Fig. 6, where correlation, PCA and LDA correspond to the results for the original images, GIC_correlation, GIC_PCA, and GIC_LDA correspond to the results for the images with GIC, and GIC_virtual_correlation, GIC_virtual_PCA, and GIC_virtual_LDA correspond to the results for the virtual images with GIC.

Fig. 6 illustrates that the recognition accuracy for the virtual images is improved greatly. When the variations due to illumination are larger, the improvement is greater. The recognition rates of correlation, PCA, and LDA on the virtual images are 87.24%, 87.99%, and 90.5%, respectively. For subset 1, subset 2, and subset 3, in which the variations due to illumination are small, the performance of three recognition methods are comparable, while in subset 4 and subset 5, LDA performs better. This indicates that the classifying ability of LDA is better than others.

In the future, we will validate the proposed method on larger face database.

5 Conclusion

This paper proposes a method to construct an LDISR using the 9 basis images under the 9 basis point light sources. The LDISR can represent almost all face images under arbitrary lighting conditions. The LDISR combined with the IRI is corresponding to the lighting conditions distinctly, and can estimate lighting conditions when given a face image and synthesize a virtual face image when given lighting conditions.

The experiments of reconstruction illustrate that the representation ability of LDISR is better than the 5-eigenimages and 3-basis images representation methods. The experiments on Yale Face Database B confirm the ability of LDISR in synthesizing a virtual face image and indicate that the virtual face images can improve greatly the accuracy of face recognition under variant lightings.

The main advantage of the proposed model is that it can be used to generate virtual images of anybody only from 9 basis face images of one person. And at the same time, the method need not know the lighting conditions or pre-calculate the IRI.

References

1 Moses Y, Adini Y, Ullman S. Face recognition: the problem of compensating for changes in illumination direction. In: Proceedings of the Third European Conference on Computer Vision. Stockholm, Sweden. Springer-Verlag, 1994, 286∼296

2 Sim T, Kanade T, Combining models and exemplars for face recognition: An illuminating example, In: Proceedings of CVPR 2001 Workshop on Models versus Exemplars in Computer Vision, Hawaii, USA. IEEE, 2001, 1∼10

3 Shashua A, Riklin-Raviv T, The quotient image: Class-based re-rendering and recognition with varying illuminations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(2): 129∼139

4 Wang H, Li Stan Z, Wang Y. Face recognition under varying lighting conditions using self quotient image. In: Proceedings of the 6th International Conference on Automatic Face and Gesture Recognition, Seoul, Korea. IEEE, 2004, 819∼824

5 Hallinan P W. A low-dimensional representation of human faces for arbitrary lighting conditions. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, USA. IEEE, 1994, 995∼999

6 Epstein R, Hallinan P W, Yuille A L. 5±2 eigenimages suffice: an empirical investigation of low dimensional lighting models. In: Proceedings of IEEE Workshop on Physics-Based Vision, Boston, USA. IEEE, 1995, 108∼116

7 Shashua A. On photometric issues in 3D visual recognition from a single 2D image. *International Journal of Computer Vision*, 1997, **21**(1):99∼122

8 Belhumeur P, Kriegman D. What is the set of images of an object under all possible illumination conditions. *International Journal of Computer Vision*, 1998, **28**(3):245∼260

9 Georghiades A S, Belhumeur P N, Kriegman D J. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, **23**(6): 643∼660

10 Lee K, Jeffrey Ho, Kriegman D. Nine points of light: Acquiring subspaces for face recognition under variable lighting. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Hawaii, USA. IEEE, 2001(1), 519∼526

11 Gao W, Shan S, Chai X, Fu X. Virtual face image generation for illumination and pose insensitive face recognition. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Hong Kong. IEEE, 2003, 776∼779

12 Zhao W, Chellappa R. Robust face recognition using symmetric shape-from-shading. Technical Report CARTR-919, 1999, Center for Automation Research, University of Maryland, College Park, MD.

13 Basri R, Jacobs D. Lambertian reflectance and linear subspaces, In: Proceedings of the 8th IEEE Computer Society International Conference On Computer Vision, Vancouver, Canada. IEEE, 2001, **2**: 383∼390

14 Ramamoorthi R, Hanrahan P. An efficient representation for irradiance environment maps. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, USA: ACM Press, 2001, 497∼500

15 Ramamoorthi R. Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(10): 1322∼1333

16 Chen H F, Belhumeur P N, Jacobs D W. In search of illumination invariants. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, South Carolina, USA. IEEE, 2000, **1**: 54∼261

17 Wang H, Li Stan Z, Wang Y, Zhang W, Illumination modeling and normalization for face recognition, In: Proceedings of IEEE International Workshop on Analysis and Modeling of Faces and Gestures, Nice, France. IEEE, 2003, 104∼111

18 Zhao J, Su Y, Wang D, Luo S. Illumination ratio image: Synthesizing and recognition with varying illuminations. *Pattern Recognition Letters*, 2003, **24**(15): 2703∼2710

19 Wen Z, Liu Z, Huang T S. Face relighting with radiance environment maps. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, USA. IEEE, 2003, **2**: 158∼165

20 Qing L, Shan S, Chen X. Face relighting for face recognition under generic illumination. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Canada. IEEE, 2004, **5**: 733∼736

HU Yuan-Kui Ph. D. candidate in the Department of Automation at University of Science and Technology of China. His research interests include face recognition, image processing, and pattern recognition.

WANG Zeng-Fu Professor of the Department of Automation at University of Science and Technology of China. His current research interests include audio & vision information processing, intelligent robots, and pattern recognition. Corresponding author of this paper. E-mail: zfwang@ustc.edu.cn