ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.18, No.9, September 2007, pp.2318−2325
DOI: 10.1360/jos182318

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

基于形态学商图像的光照归一化算法*

何晓光[1]，田 捷[1+]，毋立芳[2]，张瑶瑶[1]，杨 鑫[1]

[1](中国科学院 自动化研究所 复杂系统与智能科学实验室,北京 100080)
[2](北京工业大学 电子信息与控制工程学院,北京 100022)

Illumination Normalization with Morphological Quotient Image

HE Xiao-Guang[1], TIAN Jie[1+], WU Li-Fang[2], ZHANG Yao-Yao[1], YANG Xin[1]

[1](Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, The Chinese Academy of Sciences, Beijing 100080, China)
[2](School of Electronic Information & Control Engineering, Beijing University of Technology, Beijing 100080, China)
+ Corresponding author: Phn: +86-10-62532105, Fax: +86-10-62527995, E-mail: tian@ieee.org

He XG, Tian J, Wu LF, Zhang YY, Yang X. Illumination normalization with morphological quotient image. *Journal of Software*, 2007,18(9):2318−2325. http://www.jos.org.cn/1000-9825/18/2318.htm

Abstract: Face recognition under complex illumination conditions is still an open question. To cope with the problem, this paper proposes an effective illumination normalization method. The proposed method employs morphology and quotient image techniques by analyzing the face illumination, and it is upgraded with dynamical lighting estimation technique to strengthen illumination compensation and feature enhancement. Compared with traditional approaches, this method doesn't need any training data and any assumption on the light conditions, moreover, the enrollment requires only one image for each subject. The proposed methods are evaluated on Yale Face database B and receive a very competitive recognition rate with low computational cost.
Key words: face recognition; illumination normalization; morphology; quotient image

摘 要: 复杂光照条件下的人脸识别是一个困难但需迫切解决的问题,为此提出了一种有效的光照归一化算法.该方法根据面部光照特点,基于数学形态学和商图像技术对各种光照条件下的人脸图像进行归一化处理,并且将它发展到动态地估计光照强度,进一步增强消除光照和保留特征的效果.与传统的技术相比,该方法无须训练数据集以及假定光源位置,并且每人只需一幅注册图像.在耶鲁人脸图像库B上的测试表明,该算法以较小的计算代价取得了优良的识别性能.
关键词: 人脸识别;光照归一化;形态学;商图像
中图法分类号: TP391 **文献标识码**: A

* Supported by the Project of National Science Foundation for Distinguished Young Scholars of China under Grant No.60225008 (国家杰出青年科学基金); the Key Project of National Natural Science Foundation of China under Grant Nos.60332010, 60575007 (国家自然科学基金重点项目); the Project for Young Scientists' Foundation of National Natural Science of China under Grant No.60303022 (国家自然科学基金青年科学基金); the Project of Natural Science Foundation of Beijing of China under Grant No.4052026 (北京市自然科学基金); the Beijing Municipal Education Commission Foundation of China under Grant No.KM200610005011 (北京市教委基金)
Received 2006-10-08; Accepted 2006-12-27

1 Introduction

Face recognition benefits many fields nowadays, such as the new generation intelligent human-computer interfaces and e-services including e-home, e-shopping and e-banking. Related research activities have significantly increased over the past a few years. However, evaluations of the state-of-the-art face recognition and systems have shown that recognition performance of most current technologies degrades due to the variations of illumination, especially in the outdoor environment[1,2].

To cope with face variations caused by illumination, many techniques are developed. Generally, those approaches fall into two categories: illumination modeling and illumination normalization.

The idea of illumination modeling is to learn the extent of the variation in some suitable subspace or manifold. Recognition is then conducted by choosing the subspace or manifold closest to the novel image. This category typically includes following ones. Illumination Manifold[3] represents the object as a manifold in a low-dimensional space. Illumination Cone[4] models the illumination images distribution of a subject as a cone. Spherical Harmonic based representations[5−7] approximate the arbitrary illumination of a convex Lambertian object with a linear subspace spanned by nine harmonic images. Recently, Lee, *et al*. constructed a low dimension illumination subspace with Bilinear Illumination Model[8]. In fact, Illumination modeling methods make efforts to explore some intrinsic nature of illumination data, and are always useful for rendering and synthesizing novel faces. However, they are not so effective for face recognition. Firstly, Illumination modeling needs to learn or construct the model with a training dataset. Illumination Manifold and Illumination Cone need a large set of samples that covers illumination variation. Spherical Harmonic based representations use 3D face data[5,7] or samples under special lighting conditions[6] to obtain the harmonic images. Bilinear Illumination Model requires a 3D bootstrap dataset as well. In most applications, there are few samples for training and it is also unpractical to capture 3D data or images under controlled conditions. On the other hand, learning the suitable subspace or manifold usually involves expensive spatio-temporal cost.

Illumination normalization attempts to transform the illumination image into a certain normalized form or illumination invariant feature. Recognition is then performed by using the canonical form or invariant feature. Examples of the widely used methods are edge map, intensity derivative, Gamma correction, logarithm transform, and histogram equalization etc. However, non-uniform illumination variation, including the shadows, is still difficult to deal with by these global processing techniques. As an improvement, Adaptive Histogram Equalization[9], Region-based Histogram Equalization[10], and Block-based Histogram Equalization[11] have been proposed to conduct local normalization of face images, but the histogram equalization only enhances image pixel gray-level contrast in the spatial domain. As a new attempt, Du, *et al*.[12] used wavelet transform to enhance the contrast and the edges of face images simultaneously in the frequency domain. Discrete Cosine Transform is also employed to compensate the illumination variations in the logarithm domain[13]. Unfortunately, those methods can not meet the illumination variations caused by shadows and specularities.

Most recently, quotient image based methods[10,14−17] are reported to be a simple yet practical illumination subtraction technique and become one research direction of illumination normalization. Quotient Image (QI)[14], initially proposed by Shashua, *et al*., is defined as image ratio between a test image and linear combination of three images illuminated by non-coplanar lights. QI depends only on the albedo information, thus it is illumination free. Nevertheless, as a bootstrap database is required for this method, the performance would decline when dominant features between the bootstrap set and the test set are misaligned. Shan *et al*. proposed Quotient Illumination Relighting (QIR)[10] to recover a non-uniform lighting image with normal lighting, but the method is based on the assumption that the lighting condition of the image is known. Without the assumptions of alignment and lighting condition, H. Wang *et al*. developed Self-Quotient Image (SQI)[15,16] to estimate reflectance by the ratio of original image and its smooth version with weighted Gaussian filter. In another recent work, T. Chen *et al*. normalize the illumination image with the Total Variation based Quotient Image (TVQI)[17] and obtain a promising result. However, TVQI is very time consuming and is not competent for real-time applications.

Based on the experiential lessons of the existing illumination works, in particular on quotient image techniques, Morphological Quotient Image (MQI) and its dynamic version (DMQI) are proposed to improve the performance of face recognition under various illumination conditions in this paper. Compared with previous methods, MQI and DMQI require no training data, no assumption on the light conditions, and no alignment between different images for the illumination normalization. Moreover, they need only one enrolled image for each subject in face recognition. The results on the publicly available face database show that the proposed methods are very competitive in terms of both the recognition rate and the computational complexity.

The next section gives a detailed presentation of MQI and DMQI. Section 3 evaluates the proposed method in Yale Face Database B and compares it with other well-known approaches for the illumination problem. The last section makes a conclusion with future perspectives.

2 Morphological Quotient Image

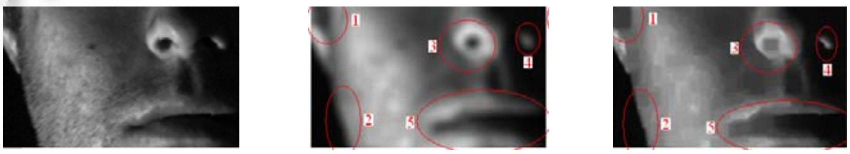
In the framework of MQI, the intensity of point (i,j) in an image I is modeled as

$$I(i,j)=\rho(i,j)l(i,j)+n(i,j) \tag{1}$$

where ρ, l, and n are components of albedo, lighting and additional noise respectively. Estimating ρ and l only from the input image I is a well-known ill-posed problem. Regarding this, some prior knowledge should be considered: (1) the reflectance ρ can be varied randomly and (2) the lighting l is smooth except edges and verges. According to the Lambertian reflectance theory, l is determined by the light source $\hat{s}$ and the surface normal $\hat{n}$ of a face: $l(i,j)=\hat{n}(i,j)\cdot\hat{s}(i,j)$. In fact, $\hat{s}$ changes slowly within the range of face surface, while $\hat{n}$ changes sharply around edges and verges, which leads l to keep smooth except edges and verges. Therefore, the lighting estimating function should satisfy the basic assumption and keep the edges and verges lossless.

Based on the model, Jobson *et al*. employ an isotropic filter[18,19] and Gross *et al*. apply an anisotropic filter[20] to get the smooth version of image I for the estimation of l. Like the traditional linear convolutions and smoothing filters, Jobson's filter creates halo effects around edge region. Gross and Brajovie's anisotropic filter can reduce the halo effect, but it is an iterative one. For real-time applications, this method is too expensive in computational cost.

In the paper, Closing operator[21], one of morphological approaches, is employed for the lighting component estimation. Closing operator is a nonlinear filter, and it is conducted with Dilation followed by Erosion. Figure 1 denotes the merits of Closeing. It tends to remove small-scale features such as narrow breaks, thin gulfs and small holes, and retain the large scale features such as edges and verges. It is less destructive of the original boundary shape of foreground.



(a) Original image (b) With traditional Gaussian filter (c) With closing operator

Fig.1 Lighting estimation

By combining with the special denoising method (in next subsection), the MQI is formally written as

$$Q = \frac{Denoise(I)}{Close(Denoise(I))} = \frac{\rho l}{l} = \rho \tag{2}$$

where *Denoise*() and *Close*() stand for the denoising method and the Closing operator respectively.

2.1 Singularity removing

Noise in the image can decrease the performance of illumination normalization and face recognition, so it is necessary to denoise the raw image before applying MQI.

There are many kinds of noises, such as salt-pepper noise, Gaussian noise and motion-blur noise etc. However, in MQI, we are only interested in singularity noise. In the paper, the so-called singularity noise is defined as the point on which the value is the unique maximum (or unique minimum) of its neighbor's values. The singularity noise is harmful for the Closing operation since the operation is just associated with local maximum and minimum. The recovering approach of singularity noise is formally modeled as

$$Denoise(I(i,j)) = \begin{cases} \underset{(n,m)\in Neighbor(i,j)}{\text{Maximum}}(I(n,m)), & \forall (n,m) \in Neighbor(i,j), I(i,j) > I(n,m) \\ \underset{(n,m)\in Neighbor(i,j)}{\text{Minimum}}(I(n,m)), & \forall (n,m) \in Neighbor(i,j), I(i,j) < I(n,m) \\ I(i,j), & \text{otherwise} \end{cases} \tag{3}$$

where *Neighbor*(i,j) stands for 8-neighbor set of point (i,j). Removing the singularities makes the Closing-based estimation of lighting more robust, and improves the performance of face recognition as denoted by the experiments in Section 3.

2.2 Dynamic morphological quotient image

The size of morphological template is a key parameter for the illumination normalization. When the size is large, MQI estimates lighting component within a large range. It can keep large-scale features well, but has poor performance of illumination compensation in local dark areas, especially in shadows. On the other side, with a small template, MQI estimates lighting component within a small range. It can get good illumination compensation in local areas, but misses the large-scale features. Therefore dynamic template is considered to enhance the illumination normalization of face image.

The scheme of Dynamic Morphological Quotient method (DMQI) is formally described as

$$DClose(i,j) = \begin{cases} Close^{l}(i,j), & Close^{l}(i,j) > \alpha \cdot Close^{s}(i,j) \\ Close^{m}(i,j), & \alpha \cdot Close^{s}(i,j) \ge Close^{l}(i,j) \ge \beta \cdot Close^{s}(i,j) \\ Close^{s}(i,j), & \beta \cdot Close^{s}(i,j) > Close^{l}(i,j) \end{cases} \tag{4}$$

where α and β are the parameters of feature scale estimation while $\alpha>\beta>1.0$. l, m and s are three different template sizes of the closing operator while $l>m>s>1$. The five parameters are determined experientially.

When $\beta \cdot Close^{s}(i,j) > Close^{l}(i,j)$, illumination estimation with the small template is almost the same as that with the large template. It demonstrates that the gray value around point (i,j) is very smooth, such as regions of cheek and forehead etc. Therefore the feature around the point is of small scale and the small template is chosen. When $Close^{l}(i,j) > \alpha \cdot Close^{s}(i,j)$, illumination estimation with the small template is greatly different from that with the large template. It denotes there are a sharp change around point (i,j), such as the areas of brow, eye, mouth, and nose. Accordingly, the feature is of large scale nearby (i,j) and the large template is chosen. DMQI introduces the technique of feature scale measurement and provides a flexible compensation for uneven illumination.

3 Experiments

In the experiments, recognition is performed with the nearest neighbor classifier. The distance is measured with normalized correlation, i.e. cosine of the angle between two image vectors. The similarity between two images I_i

and I_t is formally defined as

$$\Phi(I_i, I_t) = \cos(\angle\langle I_i, I_t\rangle) = \frac{I_i \cdot I_t}{\| I_i \| \| I_t \|} \tag{5}$$

The proposed method is evaluated in Yale Face database B[1], which includes 10 subjects and collects the frontal images of each subject under 64 different lighting conditions. Those images are divided into five subsets according to the angle the light source direction makes with the camera axis: Subset 1 (up to 12°), Subset 2 (up to 25°), Subset 3 (up to 50°), Subset 4 (up to 77°), and Subset 5 (above 77°). Seven corrupted images are discarded, and there are total 633 images: 70, 118, 118, 138, and 189 images in subset 1 to 5 respectively. The image under the frontal illumination of each subject is selected for the enrollment, and the rest face images are used for evaluation. All images are cropped to 281 pixels height by 241 pixels width, according to the eyes positions. Fig.2 shows some examples of this database.

Fig.2 Examples of Yale face database B

In the first experiment, the impacts of the singularity removing and the template size are evaluated as denoted in Table 1. With the preprocessing of singularity removing, experiments 4, 5, and 6 receive lower Recognition Error Rates (RER) than experiments 1, 2, and 3. Meanwhile, the template size is an important factor that affects the performance of MQI. As denoted by experiment 2 and 5, 7 pixels height by 7 pixels width is an optimal template size. In experiments 1 and 4, the template size is too small and only small-scale features maintain with the illumination normalization. Contrarily, in experiments 3 and 7 the template size is too large and many large-scale features are kept. Small-scale features can obtain good illumination compensation, but they have low Signal Noise Ratio (SNR) since noises are usually of small scale. Large-scale features raise high SNR, but they are not robust to uneven illumination. Therefore, the performance of experiments 1 and 4 is better than that of experiments 3 and 7 in poor illumination subsets, but they are worse than experiments 3 and 7 in good illumination subsets.

Table 1 Evaluation of MQI methods with Yale database B

ID	Preprocessing	Methods	Recognition error rate (%) vs. Illumination				
			Set 1	Set 2	Set 3	Set 4	Set 5
1	Non	MQI (size=5×5)	0.0	0.0	2.5	2.2	3.7
2	Non	MQI (size=7×7)	0.0	0.0	0.0	0.0	2.6
3	Non	MQI (size=9×9)	0.0	0.0	0.8	7.2	4.8
4	Remove singularity	MQI (size=5×5)	0.0	0.8	2.5	1.4	2.6
5	Remove singularity	MQI (size=7×7)	0.0	0.0	0.0	0.0	2.1
6	Remove singularity	MQI (size=9×9)	0.0	0.0	0.0	2.2	4.2
7	Remove singularity	DMQI (α=1.8,β=1.35)	0.0	0.0	0.0	0.0	1.1

Figure 3 shows the images of illumination normalization with the quotient image based methods. The normalized images of QI and QIR are not satisfying in the most subsets. QI and QIR learn their illumination models from a special dataset, so those models do not always fit a novel image. Over-compensation appears around some dark areas in QI, while there is a sharp contrast between bright and black regions in QIR. Over-compensation still exists on some small blocks in SQI and TVQI, especially around eyes and nose, because the texture distribution of those regions is uneven, which arises great difficulty for the local illumination estimation. However, in DMQI, those feature intensive areas are measured, and they are then normalized by a flexible compensation. There are still some pepper noises in poor illumination area. It is believed that the pepper noises are caused by underexposure, and they are propagated by the division operation[17].

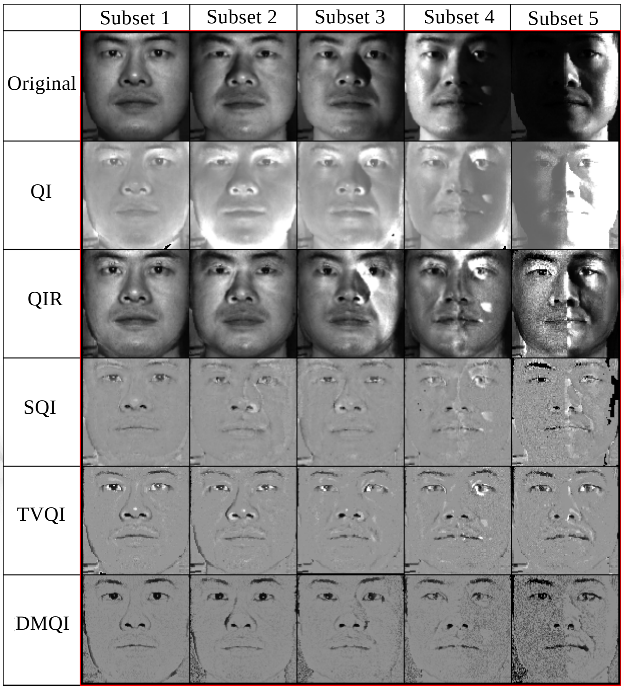


Fig.4 Illumination normalization of various quotient image based methods

Table 2 lists the RER of some well-known approaches on different illumination subsets. Many methods are not competent for Subset 5, which is collected with the worst illumination conditions, and the RER of most methods are more than 10% on that subset. SQI obtains a better result on subset 5, but it makes an over compensation of illumination and thus the performance on subset 4 is even worse than that on subset 5. The best result is achieved with TVQI, which estimates lighting image by minimizing a total variation based function. Unfortunately, TVQI need a number of iterations to find the optimal solution, and takes $O(n^{3/2})$ CPU times per iteration as denoted in Table 3.

Table 2 Recognition error rate (%) of different methods with Yale face database B

Methods	Recognition error rate (%) vs. Illumination				
	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5
Correlation[1]	N/A	0.0	23.3	73.6	N/A
Enginface w/o 1st3[1]	N/A	0.0	19.2	66.4	N/A
Linear subspace[1]	N/A	0.0	0.0	15.0	N/A
Cones cast[4]	N/A	0.0	0.0	10.0	37.0
QI[17]	0.0	1.7	38.1	65.9	76.7
QIR[17]	0.0	0.0	0.0	9.4	21.2
SQI[17]	0.0	0.0	0.0	3.6	2.1
Relative difference space[22]	0.0	0.0	0.0	0.0	12.0
MQI	0.0	0.0	0.0	0.0	2.1
DMQI	0.0	0.0	0.0	0.0	1.1
TVQI[17]	0.0	0.0	0.0	0.0	0.0

Table 3 compares MQI and DMQI with TVQI on the computational cost, where n stands for the pixel number of an image, l is the iteration time in TVQI, the average processing time is statistical calculated with the 633 images of Yale Face database B, and the computation is conducted on Dell 4700 with Pentium 4 CPU 2.80GHZ. The computational costs of MQI and DMQI are much lower than that of TVQI in terms of both time and space. Compared with TVQI, the proposed methods can be used for real-time applications.

Table 3 Computational cost comparison of MQI, DMQI, and TVQI

Items	Computational cost comparison		
	MQI	DMQI	TVQI
Average processing time per image (s)	0.08	0.20	12.89
Temporal complexity	$O(n)$	$O(3\cdot n)$	$O(l\cdot n^{3/2})$
Spatial complexity	$O(n)$	$O(n)$	$O(n\cdot \log n^{1/2})$

4 Conclusions

This paper proposes an effective yet low-cost illumination normalization method, Morphological Quotient Image (MQI), and extends it by the dynamic version (DMQI) with a flexible compensation technique. Compared with traditional approaches, the methods need no training data, no assumption on the light conditions, and no alignment between different images for illumination normalization. Moreover, they require only one image per subject for the enrollment. The proposed methods are evaluated on Yale Face database B and receive a very competitive recognition rate with low spatio-temporal cost.

Future work is focused on post processing. As shown in Fig.4, there are still some pepper-like noises in DMQI, which are caused by underexposure and propagated by the division operation of quotient image. It is a challenge task to remove the noise while keeping face feature lossless, and simultaneously it provides a further improvement space for MQI and DMQI.

Acknowledgement Special thanks for Dr. Zhang Shi-Chong, Dr. Yin Lu, and Dr. Wang Juan-Juan for the insightful discussion on some issues about this paper.

References:

[1] Georghiades A, Belhumeur P, Kriegman D. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2001,23(6):643−660.

[2] Zhao W, Chellappa R, Rosenfeld A, Phillips P. Face recognition: A literature survey. UMD CfAR Technical Report, CAR-TR-948, 2000.

[3] Murase H, Nayar S. Learning and recognition of 3D object from appearance. In: Proc. of the IEEE Workshop on Qualitative Vision. 1993. 39−50.

[4] Georghiades A, Kriegman D, Belhumeur P. Illumination cones for recognition under variable lighting: Faces. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 1998. 52−58.

[5] Basri R, Jacobs D. Lambertian reflectance and linear subspaces. In: Proc. of the Int'l Conf. on Computer Vision. 2001. 383−390.

[6] Lee K, Ho J, Kriegman D. Nine points of light: Acquiring subspaces for face recognition under variable lighting. In: Proc. of the Computer Vision & Pattern Recognition. 2001. 519−526.

[7] Zhang L, Samaras D. Face recognition under variable lighting using harmonic image exemplars. In: Proc. of the Computer Vision & Pattern Recognition. 2003. 19−25.

[8] Lee J, Moghaddam B, Pfister H, Machiraju R. A bilinear illumination model for robust face recognition. In: Proc. of the IEEE Int'l Conf. on Computer Vision. 2005. 1177−1184.

[9] Zimmerman J, Pizer S, Staab E, Perry J, McCartney W, Brenton B. An evaluation of the effectiveness of adaptive histogramequalization for contrast enhancement. IEEE Trans. on Medical Imaging, 1988,7(4):304−312.

[10] Shan S, Gao W, Cao B, Zhao D. Illumination normalization for robust face recognition against varying lighting conditions. In: Proc. of the IEEE Workshop on Analysis and Modeling of Faces and Gestures. 2003. 157−164.

[11] Xie X, Lam K. Face recognition under varying illumination based on a 2D face shape model. Pattern Recognition, 2005,38(2): 221−230.

[12] Du S, Ward R. Wavelet-Based illumination normalization for face recognition. In: Proc. of the IEEE Int'l Conf. on Image Processing. 2005. 954−957.

[13] Chen W, Er M, Wu S. Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Trans. on Systems, Man, and Cybernetics (Part B), 2006,36(2):458−466.

[14] Shashua A, Riklin-Raviv T. The quotient image: Class-Based re-rendering and recognition with varying illuminations. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2001,23(2):129−139.

[15] Wang H, Li S, Wang Y. Generalized quotient image. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition. 2004. 498−505.

[16] Wang H, Liu J, Wang Y. Self-Quotient Image. Computer Engineering, 2005,31(18):178−179,215 (in Chinese with English abstract).

[17] Chen T, Yin W, Zhou X, Comaniciu D, Huang T. Illumination normalization for face recognition and uneven background correction using total variation based image models. In: Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2005. 532−539.

[18] Jobson D, Rahman Z, Woodell G. Properties and performance of a center/surround retinex. IEEE Trans. on Image Processing, 1997, 6(3):451−462.

[19] Jobson D, Rahman Z, Woodell G. A multi-scale retinex for bridging the gap between color images and the human observation of scenes. IEEE Trans. on Image Processing, 1997,6(7):965−976.

[20] Gross R, Brajovie V. An image preprocessing algorithm for illumination invariant face recognition. In: Proc. of the Int'l Conf. on Audio and Video Based Biometric Person Authentication. 2003. 10−18.

[21] Heijmans H. Mathematical morphology: A modern approach in image processing based on algebra and geometry. SIAM Review, 1995,37(1):1−36.

[22] He X, Tian J, He Y, Yang X. Face recognition with relative difference space and SVM. In: Proc. of the Int'l Conf. of Pattern Recognition. 2006. 527−531.

附中文参考文献:

[16] 王海涛,刘俊,王阳生.自商图像.计算机工程,2005,31(18):178−179,215.

HE Xiao-Guang was born in 1983. He is a Ph.D. candidate of the Institute of Automation, the Chinese Academy of Sciences. His current research areas are pattern recognition, machine learning, image processing and their applications in biometrics.

ZHANG Yao-Yao was born in 1982. She is a master candidate of the Institute of Automation, the Chinese Academy of Sciences. Her research areas are pattern recognition, image processing and their applications in biometrics.

TIAN Jie was born in 1960. He is a professor of the Institute of Automation, the Chinese Academy of Sciences. His research areas are medical image processing and analysis, pattern recognition.

YANG Xin was born in 1972. She is an associate professor of the Institute of Automation, the Chinese Academy of Sciences. Her research areas are bioinformatics and pattern recognition.

WU Li-Fang was born in 1970. She is an associate professor of the School of Electronic Information & Control Engineering, Beijing University of Technology. Her research areas are face recognition and video analysis.