Lighting Aware Preprocessing for Face Recognition across Varying Illumination

Hu Han[1,2], Shiguang Shan[1], Laiyun Qing[2], Xilin Chen[1], and Wen Gao[1,3]

[1] Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China
[2] Graduate University of Chinese Academy of Sciences, Beijing 100049, China
[3] Institute of Digital Media, Peking University, Beijing 100871, China
{hhan,sgshan,lyqing,xlchen,wgao}@jdl.ac.cn

Abstract. Illumination variation is one of intractable yet crucial problems in face recognition and many lighting normalization approaches have been proposed in the past decades. Nevertheless, most of them preprocess all the face images in the same way thus without considering the specific lighting in each face image. In this paper, we propose a lighting aware preprocessing (LAP) method, which performs adaptive preprocessing for each testing image according to its lighting attribute. Specifically, the lighting attribute of a testing face image is first estimated by using spherical harmonic model. Then, a von Mises-Fisher (vMF) distribution learnt from a training set is exploited to model the probability that the estimated lighting belongs to normal lighting. Based on this probability, adaptive preprocessing is performed to normalize the lighting variation in the input image. Extensive experiments on Extended YaleB and Multi-PIE face databases show the effectiveness of our proposed method.

1 Introduction

Face recognition has attracted much attention in the past decades for its wide potential applications in commerce and law enforcement [1]. The challenges that a face recognition system has to face include variations in lighting, head pose, facial expression, accessory and so on. Among these factors, varying lighting conditions such as shadows, underexposure and overexposure in face imaging are intractable yet crucial problems that a practical face recognition system has to deal with. In the last decades, many approaches have been proposed to handle illumination variation problem with the goal of illumination normalization, illumination-insensitive feature extraction or illumination variation modeling. Among these approaches, many are based on image processing technique for the reason of simplicity and efficiency. In this paper, we refer these image processing based approaches as illumination preprocessing, and briefly review them in the following.

Histogram equalization (HE) [2] is one of the simplest illumination preprocessing approaches for face images, which can enhance the global contrast of one image. Logarithmic transformation (LT) [3], as a nonlinear transformation,

K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part II, LNCS 6312, pp. 308–321, 2010.
© Springer-Verlag Berlin Heidelberg 2010

tends to squeeze together the larger intensity values and stretch out the smaller ones in a face image. Jobson et al. [4] extended Retinex theory [5] to a single-scale Retinex (SSR) approach which could be used to enhance face images in improving local contrast and lightness. Based on the gamma correction technique that is widely used in Computer Graphics (CG), Shan et al. [6] proposed gamma intensity correction (GIC) in order to correct the overall brightness of a face image in accordance with a pre-defined face image with canonical lighting. Through analyzing the relationship between quotient image (QI) [7] algorithm and Retinex theory based on the reflectance-illumination model, Wang et al. [8] proposed self-quotient image (SQI) to handle the varying lighting conditions in face recognition without using a bootstrap set. Nishiyama and Yamaguchi [9] extended SQI as classified appearance-based quotient image (CAQI) in order to handle face regions with different albedo separately. Xie and Lam [10] proposed local normalization (LN) to reduce or remove the effect of uneven lighting conditions in order to get the corresponding face images under normal lighting. Considering that illumination variation mainly lies in the low-frequency band, Chen et al. [11] discarded an appropriate proportion of DCT coefficients in zigzag pattern in order to minimize the variation of face images from the same individual under different lighting conditions and then inverse DCT transform was performed to get the final illumination normalized images. Based on the reflectance-illumination imaging model, TV-L1 [12] model was introduced and analyzed in logarithm domain (LTV) by Chen et al. [13] for the purpose of decomposing a face image into large-scale and small-scale components, which correspond to illumination variation and intrinsic facial features respectively. And then only the small-scale features were used for face recognition. Xie et al. [14] reconstructed the illumination normalized face image by combining both the normalized large-scale component and smoothed small-scale component (RLS). Recently, face recognition using multi-band features are studied by Di et al. [15]. Tan and Triggs [16] presented a simple and efficient image preprocessing (PP) chain, which incorporated a series of steps such as gamma correction, Difference of Gaussian (DoG), masking and contrast equalization in order to extract illumination insensitive features for face recognition. However, most of the above approaches tend to perform illumination preprocessing equally on all the face images regardless of the particular lighting of each face image. This implies that a face image with canonical lighting will be processed like a face image with side lighting using completely the same parameter settings.

Intuitively, the above pattern that most of the existing lighting normalization approaches used to handle different lighting conditions is not optimal, since any preprocessing might bring negative effect if the input image is captured under normal lighting conditions. To reveal this possibility empirically, nine of the above-mentioned illumination preprocessing approaches, i.e., HE [2], LT [3], SSR [4], GIC [6], SQI [8], LN [10], DCT [11], LTV [13] and PP [16], are evaluated on Extended YaleB face database [17] in a traditional lighting unaware pattern. And Fisherfaces [18] is exploited as the recognition method following different illumination preprocessing approaches in our evaluation. The measurement of

the evaluation is the percentage of correcting originally-wrong matches (denoted as "positive") and reversing originally-correct matches (denoted as "negative"). The results are shown in Fig. 1, from which it is clear that most of the methods do bring some negative effects while improving the face recognition performance. Some of them may even completely counteract the positive effect, which thus limit the effectiveness of traditional lighting preprocessing approaches in improving variable lighting face recognition performance. Please note that, similar empirical observation was also reported in [19], which found that some of the preprocessing methods might result in lower recognition rates if applied to images with normal lighting.

Mathematically, most of the existing lighting normalization approaches try to have a universal method to deal with various cases. However, an image is a mapping of an object under certain lighting condition. To understand all these factors from a single image is an ill-posed problem. This is why most existing approaches reversed originally-correct matches in performing lighting normalization. In fact, it makes more sense to partition a problem as several sub-problems in handling an ill-posed problem.

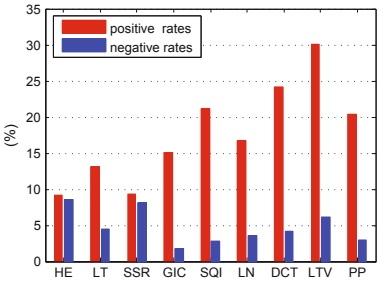

Fig. 1. The positive and negative effects of various illumination preprocessing methods performed in a lighting unaware way. We claim a "negative" if a face image is correctly recognized before the given preprocessing but incorrectly recognized after the preprocessing. On the contrary, a "positive" is reported if a face image originally incorrectly recognized can be correctly recognized after the specific preprocessing.

Based on the mathematical analysis above, we come to the idea that lighting normalization should be performed adaptively, and thus propose a lighting aware preprocessing (LAP) method for illumination-robust face recognition. Different from CAQI, in LAP, face images with different lighting conditions will be normalized in an adaptive preprocessing approach, i.e. face images with normal lighting will undergo minor or no illumination normalization, while face images with side lighting or abnormal exposure will be normalized by eliminating more large-scale components corresponding to lighting variations.

The remainder of this paper is structured as follows: Section 2 details the algorithm of the LAP and then extensive experiments are performed to verify the proposed approach in Sect. 3. Finally, we conclude this work in Sect. 4.

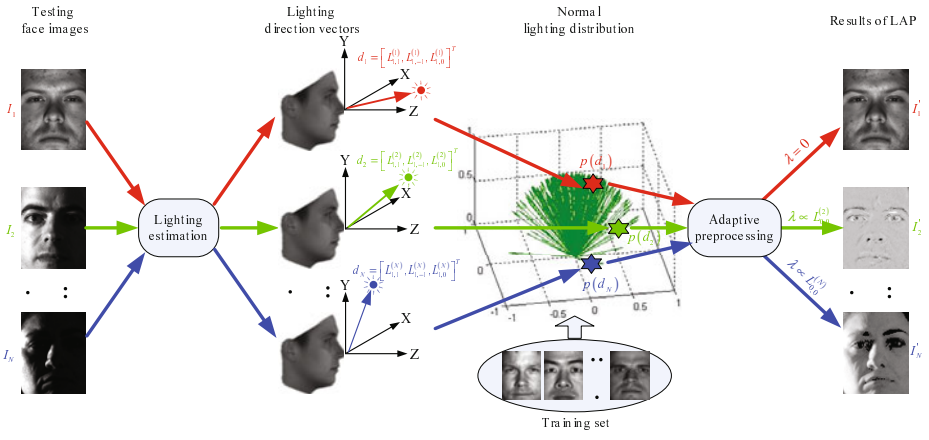

Fig. 2. Illustration for the framework of lighting aware preprocessing

2 Lighting Aware Preprocessing

In this section, we describe the details of the proposed LAP method. The algorithm overview of our LAP is shown in Fig. 2. Firstly, the lighting attribution of a testing face image is estimated by using spherical harmonic model. The estimated lighting is then analyzed by modeling the probability that it belongs to normal lighting. Finally, adaptive preprocessing is proposed to perform lighting normalization for the images with different lighting conditions. Details of each step are described below.

2.1 Lighting Attribute Estimation by Using Spherical Harmonic Model

As above mentioned, face images should be adaptively preprocessed according to their lighting conditions. Therefore, the lighting in each face image should be estimated. With different constraints introduced, many approaches have been proposed to recover the lighting from a single input face image, such as shape from shading (SFS) [20], 3D subspaces [21], 5D subspace [22], 9D linear subspace[23], illumination cone [17,24] and so on. In our LAP approach, lighting attribute is estimated by using spherical harmonic model, which has been used to estimate the harmonic basis face images that span a linear subspace to approximate a wide variety of illumination variations [23,25,26,27,28,29].

By simplifying the face imaging procedure as a convex Lambertian object under distant isotropic illumination, the image intensity is proportional to the radiance reflected by the face surface and can be approximated by

$$I(x, y) \approx \lambda(x, y) E(\alpha(x, y), \beta(x, y)) \tag{1}$$

where (x, y) ranges over the whole face surface, $\lambda(x, y)$ is the albedo at point (x, y), (α, β) is the normal at point (x, y) and $E(\alpha, \beta)$ is the total irradiance that arrives at point (x, y), which is a function of the surface normal (α, β) [30]

$$E(\alpha, \beta) = \int_{\varphi_i=0}^{2\pi} \int_{\theta_i=0}^{\pi/2} L_i((x, y), \varphi_i, \theta_i) \cos\theta_i \sin\theta_i d\theta_i d\varphi_i \tag{2}$$

where θ_i and φ_i are respectively the elevation and azimuth angles of incident light. Under the distant illumination assumption, $E(\alpha, \beta)$ is independent of surface position (x, y) [25]

$$E(\alpha, \beta) = \int_{\varphi_i=0}^{2\pi} \int_{\theta_i=0}^{\pi/2} L_i(\varphi_i, \theta_i) \cos\theta_i \sin\theta_i d\theta_i d\varphi_i \tag{3}$$

where $L_i(\varphi_i, \theta_i)$ is the radiance of the incident light with direction (φ_i, θ_i). Hence, lighting estimation is converted to recovering the coefficients $L_i(\varphi_i, \theta_i)$ given an input face image I.

As is shown independently by Basri and Jacobs [23] as well as Ramamoorthi and Hanrahan [25], $E(\alpha, \beta)$ can be well approximated by a combination of the first nine spherical harmonics

$$E(\alpha, \beta) = \sum_{l=0}^{2} \sum_{m=-l}^{l} (\frac{4\pi}{2l+1})^{1/2} A_l L_{l,m} Y_{l,m}(\alpha, \beta) \tag{4}$$

where A_l is the spherical harmonic coefficient for transfer function, $L_{l,m}$ is the coefficient of incident lighting and $Y_{l,m}$ forms the orthonormal spherical harmonic basis. It is more convenient to parameterize $Y_{l,m}$ in Cartesian coordinate system as below [23]

$$\begin{aligned}
&Y_{0,0} = \sqrt{\tfrac{1}{4\pi}} && Y_{1,-1} = \sqrt{\tfrac{3}{4\pi}} y \\
&Y_{1,0} = \sqrt{\tfrac{3}{4\pi}} z && Y_{1,1} = \sqrt{\tfrac{3}{4\pi}} x \\
&Y_{2,-2} = \sqrt{\tfrac{15}{4\pi}} xy && Y_{2,-1} = \sqrt{\tfrac{15}{4\pi}} yz \\
&Y_{2,0} = \sqrt{\tfrac{5}{16\pi}} (3z^2 - 1) && Y_{2,1} = \sqrt{\tfrac{15}{4\pi}} zx \\
&Y_{2,2} = \sqrt{\tfrac{5}{16\pi}} (x^2 - y^2)
\end{aligned} \tag{5}$$

where (x, y, z) is the representation for surface normal (α, β) in Cartesian coordinate system. Combining (1) with (4), we will get

$$\begin{aligned}
I(x, y) &\approx \sum_{l=0}^{2} \sum_{m=-l}^{l} (\tfrac{4\pi}{2l+1})^{1/2} \lambda(x, y) A_l L_{l,m} Y_{l,m}(\alpha(x, y), \beta(x, y)) \\
&= \sum_{l=0}^{2} \sum_{m=-l}^{l} L_{l,m} b_{l,m}(x, y)
\end{aligned} \tag{6}$$

where $b_{l,m}(x, y)$ is the harmonic image of a face

$$b_{l,m}(x, y) = (\frac{4\pi}{2l+1})^{1/2} \lambda(x, y) A_l Y_{l,m}(\alpha(x, y), \beta(x, y)) \tag{7}$$

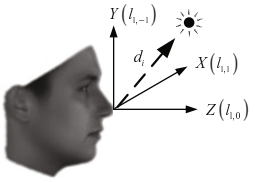

Fig. 3. The construction of a lighting direction vector

In order to estimate the 9 illumination coefficients $L_{l,m}$, one needs to know the albedo map $\lambda(x, y)$ and normal map $(\alpha(x, y), \beta(x, y))$ of the given face. However, in practice, they are usually unavailable for a single input face image. Fortunately, as shown in [27], with a quasi-constant albedo map and a warped generic 3D facial normal map as the approximations for the real ones, the 9 illumination coefficients $L_{l,m}$ can be well estimated by solving the following least squares problem

$$\hat{L} = \arg\min_{L} \|I - BL\|_{L_2} \tag{8}$$

where image I is vectorized as a P-dimensional column vector, B is a $P \times 9$ matrix with $b_{l,m}$ as its columns.

In our implementation, given an input face image, its two eyes are first localized and used to roughly align a generic 3D facial normal map. Then, spherical harmonic images, i.e. B, of this face are computed based on (7). And finally the 9 coefficients are estimated by solving (8).

According to the spherical harmonics theory, among the 9 illumination coefficients, $L_{0,0}$ is the DC component reflecting the average energy of the incident lighting, while the three first-order coefficients, $L_{1,1}, L_{1,-1}, L_{1,0}$, as illustrated in Fig. 3, reflect the intensity of incident lights in X, Y, Z directions respectively. Therefore, they are utilized in our method to form the lighting direction vector $d^{'} = [L_{1,1}, L_{1,-1}, L_{1,0}]^T$. Since we care only the relative quantity of these coefficients, we further normalize it by dividing its module and thus get a unit vector in L_2 norm $d = d^{'}/\|d^{'}\| = [l_{1,1}, l_{1,-1}, l_{1,0}]^T$, which is then used to analyze the lighting condition of the input image in the following.

2.2 Lighting Analysis with vMF Model

With the above estimated lighting attribute, what we need to do next is determining which kind of lighting it belongs to. However, it is difficult to make a quantitative definition of lighting category, as lighting condition is a subjective concept. To overcome the uncertainty of imaging procedure and the subjectiveness in lighting condition definition, we apply a statistical model to determine the probability that the estimated lighting belongs to normal lighting. The statistical model, which combines the principle of physics, geometrical model and the robustness of statistics, thus provides a relative definition of different lighting conditions instead of an absolute one.

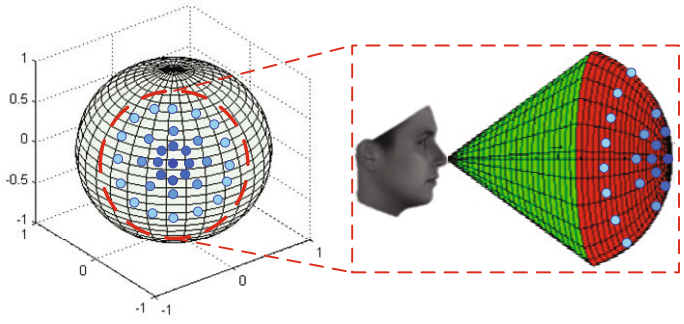

Fig. 4. The distribution of normal lighting is analogous to a Gaussian distribution on a sphere surface

In most face recognition testing protocols [17,31], so-called normal lighting usually means frontal distant lighting. Therefore, the subset with normal lighting in the testing protocol of each face database is utilized to learn a statistical model for normal lighting. In practice, normal lighting should distribute analogously to a Gaussian distribution on a unit sphere as illustrated in Fig. 4 with $d_0 = [0, 0, 1]^T$ being the expectation. Thus, normal lighting can be modeled as a von Mises-Fisher (vMF) distribution [32] which is widely used in directional statistics.

Specifically, a 3-dimensional unit random vector x(i.e., $x \in \mathbb{R}^3$ and $\|x\| = 1$) is of 3-variate von Mises-Fisher distribution if its probability density function is with the form

$$p(x|\mu, \kappa) = c(\kappa) \exp(\kappa \mu^T x) \tag{9}$$

where μ is the mean direction with $\|x\| = 1$, $\kappa(\kappa \geq 0)$ is the concentration parameter describing how strongly the unit random vectors sampled from the distribution are concentrated toward the mean direction, and normalization constant $c(\kappa)$ is defined as

$$c(\kappa) = \frac{\kappa}{4\pi \sinh \kappa} = \frac{\kappa}{2\pi(e^{\kappa} - e^{-\kappa})} \tag{10}$$

Given the vMF model, modeling normal lighting is then to estimate the parameters of the vMF model. In this study, maximum likelihood estimation is adopted to estimate μ and κ from a learning dataset. Formally, given a training set containing N face images captured under "normal" lighting conditions, we estimate the lighting direction vectors by the method in Section 2.1 for all the training images and obtain

$$D = \{d_i \in \mathbb{R}^3, 1 \leq i \leq N\} \tag{11}$$

By safely assuming d_i to be independent with each other, we have the following likelihood

$$p(D|\mu, \kappa) = \prod_{i=1}^{N} p(d_i|\mu, \kappa) \tag{12}$$

And then the log-likelihood will be

$$\ln p(D|\mu,\kappa) = N \ln c(\kappa) + \kappa \mu^T t \tag{13}$$

where $t = \sum_{i=1}^{N} d_i$. In order to get the maximum likelihood estimates for μ and κ, Lagrange multipliers is used to maximize the log-likelihood objective function

$$\Lambda(\mu,\kappa,d_i,\lambda) = N \ln c(\kappa) + \kappa \mu^T t + \lambda(1-\mu^T\mu) \tag{14}$$

subject to the constraint $\mu^T\mu = 1(\|\mu\| = 1)$. Let the derivative $d\Lambda = 0$, then we get the following system of equations

$$\begin{array}{l} \frac{\partial \Lambda}{\partial \mu} = \kappa t - 2\lambda\mu \quad\quad = 0 \\ \frac{\partial \Lambda}{\partial \kappa} = \frac{N c^{'}(\kappa)}{c(\kappa)} + \mu^T t = 0 \\ \frac{\partial \Lambda}{\partial \lambda} = 1 - \mu^T\mu \quad\quad = 0 \end{array} \tag{15}$$

From (15), it is not difficult to get an estimate for μ

$$\hat{\mu} = \frac{t}{\|t\|} \tag{16}$$

In directional statistics, the concentration parameter κ is usually estimated in an approximation manner [32,33] and for a 3-variate von Mises-Fisher distribution, the following approximation will be sufficient

$$\hat{\kappa} = \frac{3\overline{t} - \overline{t}^3}{1-\overline{t}^2} \tag{17}$$

where $\overline{t} = \|t\|/N$

After μ and κ are estimated, the statistical model for describing normal lighting is constructed and then the probability that the estimated lighting d of a testing face image belongs to normal lighting can be calculated based on (9)

$$p(d|\hat{\mu},\hat{\kappa}) = c(\hat{\kappa}) \exp(\hat{\kappa}\hat{\mu}^T x) \tag{18}$$

In this paper, the subset#1 from Extended YaleB face database is used as the training set. The face images in subset#1 are captured with the angle between the light source direction and the camera axis within 12°. Details about the subset division for Extended YaleB can be found in [17].

2.3 Adaptive Lighting Preprocessing

As we have mentioned before, once the lighting condition of a testing face image has been grouped in a relative manner, facial images will be handled accordingly. For this purpose, we further propose an adaptive method to perform illumination normalization for each testing face image. By varying the truncation scale, many existing approaches, e.g. the Gaussian smoothing filter used in [4,8], the DCT

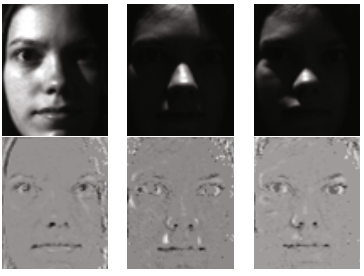

Fig. 5. The results of traditional LTV on three face images of one individual

reported in [11] and the TV-L1 model [12] utilized in LTV [13], can reach a better balance between eliminating extrinsic lighting variation and preserving intrinsic facial features. Without loss of generality, here TV-L1 model is used to implement adaptive preprocessing based on the estimated probability that the lighting in a testing face image belongs to normal lighting.

TV-L1 model aims at decomposing a face image into large-scale component u which corresponds to illumination variation and small-scale component v which corresponds to intrinsic facial feature and the large-scale component in a face image is estimated by solving the following variational problem

$$\hat{\mu} = \arg\min_{u} \int |\nabla u| + \lambda \|I - u\|_{L_1} \tag{19}$$

where $\int |\nabla u|$ is the total variation of u and λ is a scalar constant controlling the scale truncation. With u solved, the small-scale component v can be calculated as $v = I - u$, which can then be used for face recognition across varying lighting conditions. Evidently, in TV-L1 model, the scale-truncation constant λ actually balances the illumination removal by u and feature preserving in v. However, in LTV, it is empirically set and kept the same for all face images. This might be questionable, since different lighting attributes imply illumination component of different scales. Figure 5 shows some examples of LTV with fixed λ for images of the same person but with different lighting attributes. It is clear that the results are not desirable. Different from LTV, in our adaptive lighting preprocessing, TV-L1 model are applied in an adaptive pattern based on the above estimated probability rather than in a fixed pattern.

According to the analysis for parameter λ in [12], a larger truncation scale is more desirable in order to avoid discarding too much intrinsic facial features for face images with normal lighting and correspondingly a smaller λ should be used for TV-L1 model. While the effect introduced by abnormal lighting, such as the artificial edges caused by side lighting, mainly lies in high frequency band; therefore a small truncation scale is suitable and correspondingly a larger λ should be taken.

According to the above analysis, the parameter λ in TV-L1 model can be approximately determined based on the probability that the lighting in a face image belongs to normal lighting

$$\lambda = (1 - p(d|\hat{\mu}, \hat{\kappa}))\beta \tag{20}$$

where $p(d|\hat{\mu}, \hat{\kappa})$ is the above estimated probability that the lighting in a testing face image belongs to normal lighting, β is the range for parameter λ. In TV-L1 model, parameter λ can be set as any positive real number, but in practice, for face image with the size of 64×80, λ in the range of $[0, 1.2]$ will be sufficient for handling most of the lighting variations. A linear relationship between p and λ seems simple but reveals to be effective in our experiments. To be note that in the theory of TV-L1, features of all scales should be keep in v when $\lambda = 0$, i.e. $v_{\lambda=0} = I$; however, due to the limitation in computation, v cannot be calculated when $\lambda = 0$. Therefore, we force $v_{\lambda=0} = I$ in our implementation. When TV-L1 model is substituted by other methods, e.g. Gaussian smoothing filter or DCT, the parameters can also be determined like in (20). All the gallery images are also preprocessed in the same way as each testing face image when performing face recognition.

3 Experimental Results

3.1 Databases and Settings for Experiments

Extended YaleB [17], PIE [34] and Multi-PIE [31] are three representative face databases in the area, however, many illumination preprocessing approaches, including the proposed LAP, have gotten 100% recognition performance. Therefore, two challenging face databases of the three: Extended YaleB [17] and Multi-PIE [34], are exploited in our experiments to compare our proposed approach with other illumination preprocessing approaches in face recognition across varying illumination.

Extended YaleB face database includes the original YaleB face database with 10 individuals under 64 different illumination conditions and the extended part with 28 individuals that are also captured under 64 different illumination conditions. Totally 2,432 face images of 38 individuals under 64 illumination conditions in frontal view are used for experiments. All the face images are divided into five subsets according to [17], in which subset#1 is used as the training set for both lighting estimation and face recognition algorithm. The varying lighting in Extended YaleB is harsh for illumination-robust recognition as the lighting directions vary from left 130° degrees to right 130°.

Multi-PIE is a recently published face database, which contains as many as 755,370 images from 337 subjects, imaged under 15 view points and 19 illumination conditions in up to four recording sessions [31]. According to the testing protocol in [31], the face images of 14 randomly selected subjects are used for training and images of all the other 323 subjects are used for testing. Among all the testing images, only one face image of each individual recorded without flashes is used as gallery. The huge database size and time span of Multi-PIE have determined the challenge for variable lighting face recognition. Moreover, the limitation of 14 subjects for training further increase the difficulty in recognition across varying lighting conditions.

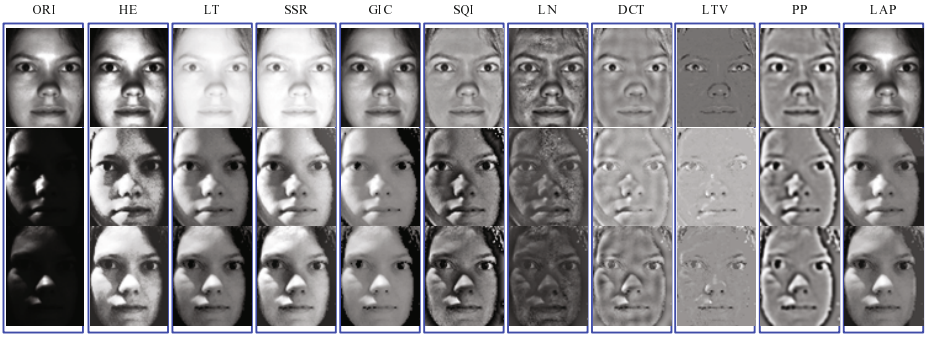

Fig. 6. Illumination preprocessing on testing face images using different approaches. Images in the first column are the original input face images under different lighting conditions. Images in the rest columns are the results of different illumination preprocessing approaches.

Before performing any illumination preprocessing, all the face images are geometrically normalized to the size of 64×80 with the distance between two eyes 35 pixels. The proposed LAP is a kind of illumination normalization approach instead of illumination-insensitive feature extraction or illumination variation modeling approach, therefore, the state of the art as well as several representative illumination normalization approaches are taken for comparison, i.e., HE [2], LT [3], SSR [4], GIC [6], SQI [8], LN [10], DCT [11], LTV [13] and PP [16]. For fair comparison with other methods, we exploited the parameters settings recommended in the original literature proposing the corresponding methods for comparison. As what we concern is the comparison between different lighting preprocessing approaches, Fisherfaces [18] is fixed as the recognition algorithm for all the illumination preprocessing approaches we compared.

Face recognition is performed on the illumination normalized face images preprocessed by different approaches and recognition performance is reported to verify the effectiveness of different lighting preprocessing approaches in improving the robustness for face recognition across varying lighting conditions. For the convenience of our description, we denote "ORI" as the original face images without any lighting preprocessing.

3.2 Comparisons

The visualization of some illumination normalized face image of different lighting preprocessing approaches is illustrated in Fig. 6. In the figure, the face images in the first column are the original input images and those in the rest columns are the results of different lighting preprocessing approaches labeled above the column. As can be seen from the figure, the traditional approaches performed in a lighting unaware pattern tend to produce satisfying results for some kinds of lighting but not for others. On the contrary, for our LAP, testing images with normal lighting are kept as close to the original as possible and images

Table 1. Face recognition performance of Fisherfaces on the preprocessed face images by different illumination preprocessing approaches from Extended YaleB and Multi-PIE databases

Approach	Recognition Rate (%)	
	Extended YaleB	Multi-PIE
ORI	54.15	52.77
HE	54.75	62.53
LT	62.79	62.73
SSR	55.45	63.79
GIC	67.73	64.27
SQI	72.58	65.71
LN	67.36	60.74
LDCT	74.10	61.82
LTV	78.02	60.81
PP	71.56	61.18
LAP	**86.89**	**71.15**

with abnormal exposure are processed to discard most lighting variations while preserving more discriminative facial features compared with LTV.

Face recognition experiments are then performed on the two face databases following different illumination preprocessing approaches and the recognition rates are reported in Table 1. As shown in the table, our LAP achieves impressively better face recognition performance than all the other methods on both Extended YaleB and Multi-PIE face databases. On Extended YaleB, compared with LTV, LAP gets more than 8% higher face recognition rate. Even on the much more challenging Multi-PIE face database, LAP gets the highest face recognition rate 71.15%. Experimental results on Extended YaleB and Multi-PIE face databases suggest that our proposed LAP framework is more effective and robust in improving face recognition performance across varying illumination compared with the traditional lighting unaware approaches.

4 Conclusions

Traditional illumination preprocessing methods deal with face images in a lighting unaware way, so they might suffer from negative effect, for instance, failing to recognize an image which can correctly recognized before preprocessing. This paper analyzed the problem and proposes a lighting aware preprocessing method. In the method, face images with different lighting conditions are processed according to the lighting attribute in the images. Experiments illustrate impressive performance improvement compared with the state of the art and representative illumination preprocessing methods. To be note that although TV-L1 is utilized in the proposed LAP framework, other methods such as low-pass filtering and DCT can also be embedded into the proposed LAP framework.

The preliminary studies in this paper show that we still have large space for improvement for illumination-invariant face recognition. Preprocessing automatically adapted to the lighting attribute of the image might be a promising possibility.

Currently, spherical harmonic model is used to estimate the lighting in a testing face image. Simple and efficient approaches without using 3D face information, e.g. the method proposed by S. Choi, et al. [35], might be used for lighting estimation. Moreover, the relationship between the normal lighting probability and adaptive parameter selection will also be exploited in future work.

Acknowledgments

This paper is partially supported by Natural Science Foundation of China under contracts No.60803084, No.60872077, and No. U0835005; National Basic Research Program of China (973 Program) under contract 2009CB320902, and ISVISION Technology Co. Ltd.

References

1. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.J.: Face recognition: A literature survey. ACM Computing Surveys 35, 399–458 (2003)
2. Gonzalez, R., Woods, R.: Digital image processing, pp. 91–94. Prentice Hall, USA (1992)
3. Adini, Y., Moses, Y., Ullman, S.: Face recognition: The problem of compensating for changes in illumination direction. IEEE Trans. PAMI 19, 721–732 (1997)
4. Jobson, D.J., Rahman, Z., Woodell, G.A.: Properties and performance of a center/surround retinex. IEEE Trans. IP 6, 451–462 (1997)
5. Land, E.H.: An alternative technique for the computation of the designator in the retinex theory of color vision. Proc. Nati. Acad. Sci. USA 83, 3078–3080 (1986)
6. Shan, S., Gao, W., Cao, B., Zhao, D.: Illumination normalization for robust face recognition against varying lighting conditions. In: Proc. AMFG, Nice, pp. 157–164 (2003)
7. Shashua, A., Raviv, T.R.: The quotient image: Class-based re-rendering and recognition with varying illuminations. IEEE Trans. PAMI 23, 129–139 (2001)
8. Wang, H., Li, S., Wang, Y.: Face recognition under varying lighting conditions using self quotient image. In: Proc. FG, Seoul, pp. 819–824 (2004)
9. Nishiyama, M., Yamaguchi, O.: Face recognition using the classified appearanceebased quotient image. In: Proc. FG, Southampton, pp. 49–54 (2006)
10. Xie, X., Lam, K.: An efficient illumination normalization method for face recognition. Pattern Recognition Letters 27, 609–617 (2006)
11. Chen, W., Er, M.J., Wu, S.: Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Trans. SMC:B 36, 458–466 (2006)
12. Chan, T., Esedoglu, S.: Aspects of total variation regularized l1 function approximation. CAM Report, 4–7 (2004)
13. Chen, T., Yin, W., Zhou, X.S., Comaniciu, D., Huang, T.S.: Total variation models for variable lighting face recognition. IEEE Trans. PAMI 28, 1519–1524 (2006)
14. Xie, X., Zheng, W., Lai, J., Yuen, P.C.: Face illumination normalization on large and small scale features. In: Proc. CVPR, Alaska, pp. 1–8 (2008)
15. Di, W., Zhang, L., Zhang, D., Pan, Q.: Studies on hyperspectral face recognition with feature band selection. IEEE Trans. SMC-A (to appear)

16. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. In: Proc. ICCV Workshop, Rio de Janeiro, pp. 168–182 (2007)
17. 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 Trans. PAMI 23, 643–660 (2001)
18. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Trans. PAMI 19, 711–720 (1997)
19. Du, B., Shan, S., Qing, L., Gao, W.: Empirical comparisons of several preprocessing methods for illumination insensitive face recognition. In: Proc. ICASSP, Pennsylvania, pp. 981–984 (2005)
20. Horn, B.K.P., Brooks, M.J.: The variational approach to shape from shading. CVGIP 33, 174–208 (1986)
21. Shashua, A.: On photometric issues in 3d visual recognition from a single 2d image. IJCV 21, 99–122 (1997)
22. Hallinan, P.W.: A low-dimensional representation of human faces for arbitrary lighting conditions. In: Proc. CVPR, Seattle, pp. 995–999 (1994)
23. Basri, R., Jacobs, D.W.: Lambertian reectance and linear subspaces. IEEE Trans. PAMI 25, 218–233 (2003)
24. Belhumeur, P.N., Kriegman, D.J.: What is the set of images of an object under all possible illumination conditions? IJCV 28, 245–260 (1998)
25. Ramamoorthi, R., Hanrahan, P.: On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. JOSA 18, 2448–2459 (2001)
26. Zhang, L., Samaras, D.: Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics. IEEE Trans. PAMI 28, 351–363 (2006)
27. Qing, L., Shan, S., Gao, W., Du, B.: Face recognition under generic illumination based on harmonic relighting. IJPRAI 19, 513–531 (2005)
28. Wang, Y., Liu, Z., Hua, G., Wen, Z., Zhang, Z., Samaras, D.: Face re-lighting from a single image under harsh lighting conditions. In: Proc. CVPR, Minnesota, pp. 1–8 (2007)
29. Jiang, X., Kong, Y.O., Huang, J., Zhao, R.-c., Zhang, Y.: Learning from real images to model lighting variations for face images. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 284–297. Springer, Heidelberg (2008)
30. Forsyth, D.A., Ponce, J.: Computer vision: A modern approach, pp. 46–58. Prentice Hall, USA (2002)
31. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-pie. Image and Vision Computing 28, 807–813 (2010)
32. Mardia, K.V., Jupp, P.E.: Directional statistics, pp. 36–44. J. Wiley, Chichester (2000)
33. Banerjee, A., Dhillon, I.S., Ghosh, J., Sra, S.: Clustering on the unit hypersphere using von mises-fisher distributions. JMLR 9, 1345–1382 (2005)
34. Sim, T., Baker, S., Bsat, M.: The cmu pose, illumination, and expression database. IEEE Trans. PAMI 25, 1615–1618 (2003)
35. Choi, S., Kim, C., Choi, C.: Shadow compensation in 2d images for face recognition. Pattern Recognition 40, 2118–2125 (2007)