Logarithm Gradient Histogram: A General Illumination Invariant Descriptor for Face Recognition

Jun-Yong Zhu Wei-Shi Zheng* Jian-Huang Lai

Abstract— **In the last decade, illumination problem has been the bottleneck of robust face recognition system. Extracting illumination invariant features becomes more and more significant to solve this issue. However, existing works in this field only consider the variations caused by lighting direction or magnitude (denoted as *homogeneous lighting*), while the spectral wavelength is always ignored in most of the developed illumination invariant descriptors. In this paper, we claim that the spectral wavelength is important, and we propose a novel gradient based descriptor, namely Logarithm Gradient Histogram (LGH), which takes the illumination direction, magnitude and even the spectral wavelength together into consideration (denoted as *heterogeneous lighting*). Our proposal contributes in the following three-folds: (1) we incorporate homogeneous filtering to alleviate the illumination effect for each image and extract two illumination invariant components, namely logarithm gradient orientation (LGO) and logarithm gradient magnitude (LGM); (2) we propose an effective postprocessing strategy to guarantee the fault-tolerant ability and generate a histogram representation to integrate both LGO and LGM; (3) we present thorough theoretical analysis on the illumination invariant properties for our proposed method. Experimental results on CMU-PIE, Extended YaleB and HFB databases are reported to verify the effectiveness of our proposed method.**

I. Introduction

The illumination problem, as a challenging issue in the face recognition, has become a barrier in the development of many face related applications, such as video surveillance, face detection, cooperative user applications, etc. The well known face recognition vendor test (FRVT) 2006 [1] have also revealed that large variation in illumination would probably affect the performance of face recognition algorithms. A variety of works have been proposed to address this issue and they mainly fall into three categories [2]: preprocessing and normalization techniques [3], [4], face modeling based approaches [5], [6] and invariant feature extraction [7]–[12].

Preprocessing and normalization methods like histogram equalization (HE) [3] attempt to normalize face images using image processing techniques such that the processed images appear to be consistent under different lighting conditions. However, these methods are hard to obtain notable improvement in recognition though the visual effects appeared acceptable. To further investigate the cause of illumination problem, the modeling based approaches turn to explore the mechanism of face imaging. Based on the assumption that the surface of face is Lambertian, images of the same face under varying lighting conditions span a low dimensional linear subspace [5], [13]. In theory, these methods describe the illumination variation quite well, but they need a great deal of training samples to learn the variation and easily suffer from the over-fitting problem, which largely restricts their use in real applications. Compared to the above two categories, invariant feature based methods are more effective and do not demand learning. Classical methods such as Local Binary Pattern (LBP) [7] and Gabor [14] are commonly believed to be robust to slight illumination change, but their performance will drop when the lighting condition becomes severe. To overcome this problem, a variety of state-of-the-art methods have been proposed by extracting the reflectance component [2], [8], [9] or alleviating the illumination component based on the Lambert's reflectance model [11], [12]. Great success using these effective methods has been seen.

The illumination problem mentioned above is mainly referred to the variation caused by either varying lighting direction or varying lighting magnitude. In those methods, a main assumption is made that the wavelengths of light are the same, which is denoted as *homogeneous lighting* in this paper. However, it cannot hold in realistic applications. For example, the lighting wavelengths of indoor and outdoor conditions are always different, so as to the case of visible (VIS) and near infrared (NIR) spectral face images [15]. As a result, the reflectance component, which is determined by the albedo and the normal direction of facial surface, will change with the varying wavelength since the albedo is related to the spectral wavelength. For convenient, we denote the lighting condition with different spectral wavelengths as *heterogeneous lighting*. As far as we know, there is still lack of work that addresses this issue for solving the illumination problem and makes theoretical development on an valid image descriptor in this aspect. Some related works like [15] and [16] consider the VIS-NIR face matching as a multi-modality face recognition problem rather than an illumination related task. In this paper, inspired by the gradient face [11], we propose a novel gradient based descriptor, namely *logarithm gradient histogram* (LGH), and provide an in-depth analysis on its illumination invariant property. Experimental results on three public face databases verify the effectiveness of our proposed method for the illumination problem.

The rest of this paper is organized as follows: Section II further discusses the invariant feature extraction approaches.

*Corresponding author

Jun-Yong Zhu is with School of Mathematics and Computational Science, Sun Yat-sen University, China jonesjunyong@ieee.org

Wei-Shi Zheng is with the School of Information Science and Technology, Sun Yat-sen University, China wszheng@ieee.org

Jian-Huang Lai is with the School of Information Science and Technology, Sun Yat-sen University, China stsljh@mail.sysu.edu.cn

Section III formulates our descriptor by introducing the homomorphic filtering and two newly proposed illumination invariant components. Meanwhile, the theoretical proof of illumination invariant properties of our proposed method is given in this part. Experiments on CMU-PIE, Extended YaleB and HFB databases are carried out in Section IV to evaluate the performance of our proposed LGH. Finally, the conclusion of this paper is drawn in Section V.

II. Related Work

According to the Lambertian Law, the intensity of the illuminated image I can be formulated as $I(x, y) = F(x, y)L(x, y)$, that is, a product of the illumination component $L(x, y)$ and the reflectance component $F(x, y)$. As commonly assumed, the $L(x, y)$ in the Lambert's reflectance model changes very slowly and $F(x, y)$ is independent of lighting condition [11], [17], [18]. Thus, $F(x, y)$ is commonly regarded as illumination invariant feature.

Hence, in order to extract the illumination insensitive component only related to $F(x, y)$, Self Quotient Image (SQI) [8] alleviates the effect of illumination by dividing itself with the blurred version. Logarithmic Total Variation (LTV) [9] incorporates TV model to preserve the edge information and obtain a more elaborate representation. Following a similar way, Xie et al. [2] computed a better reflectance component by using Logarithmic Non-subsampled Contourlet Transform and obtained significant improvement at the cost of time consuming.

There are also some other works investigating the illumination invariant property by considering the normalized local intensity contrast, such as relative gradient [19] and weber face (WF) [12]. Inspired by the Weber's Law [20], the authors in [12] show that the radio between local difference and the center degree is insensitive to the illumination change and encouraging results are obtained. Most of the aforementioned methods verify their robustness against varying lighting conditions but few of them provide theoretical proofs for the illumination invariant properties. One exception is that Zhang et al. proved that the gradient orientation, denoted as Gradient Face (GF) [11], is somehow insensitive to the illumination influence both in theoretical analysis and experimental validation. Though great success has been achieved in [11], there are still some limitations. First, the gradient orientations of all pixels are involved in the pixel-wise comparison, taking the noise and face unrelated information into account, which will probably degenerate the recognition performance. Meanwhile, some important information in gradient domain like the gradient magnitude is neglected since it is not supported by the illumination invariant property. Also, it does not discuss the variation caused by the lighting wavelength. However, all these issues will be solved in our new proposed method.

Most popular models assume that the face images captured under various lighting conditions share the same spectral wavelength, i.e., the reflectance component $F(x, y)$ is independent of lighting variance $L(x, y)$. However, as previously mentioned, the reflectance component should be affected by lighting wavelength in reality. Hence the extracted reflectance components of the same subject under heterogeneous lighting such as sunlight, electric lamp, near-infrared camera, etc, are different from each other. To tackle this problem, [15] suggests encoding the local pattern via LBP and followed by DoG filtering, and [16] applies SIFT [21] descriptor instead. However, these two methods do not provide theoretical analysis on the illumination invariant property and they could only be suitable for dealing with VIS-NIR face matching.

In this paper, we study a more general case for the illumination problem where the lighting direction, magnitude, as well as the spectral wavelength will change in different lighting conditions. We inherit the high efficiency property of feature extraction approaches and propose an illumination invariant descriptor based on the gradient information in the logarithm domain. The main contributions of this paper are summarized as follow:

1) We introduce homogeneous filtering as a preprocessing to constrain the illumination effect and enhance facial information.
2) We propose a new histogram based illumination invariant feature descriptor LGH by integrating both gradient magnitude and gradient orientation in the logarithm domain.
3) Thorough theoretical analysis on the illumination invariant property is given for our proposed LGH, guaranteeing its performance in both homogeneous and heterogeneous lightings.
4) Experimental results show that our proposed method outperforms the related state-of-the-art approaches in the field of illumination problem.

III. Logarithm Gradient Histogram

In this section, we are going to elaborate our proposed LGH in three folds: (i) the homomorphic filtering used for constraining the illumination effect and enhancing facial information; (ii) two illumination invariant components, i.e., logarithm gradient orientation (LGO) and logarithm gradient magnitude (LGM), and the theoretical analysis on the illumination invariant property; (iii) post-precessing for integrating LGO and enhanced LGM into the histogram based feature representation.

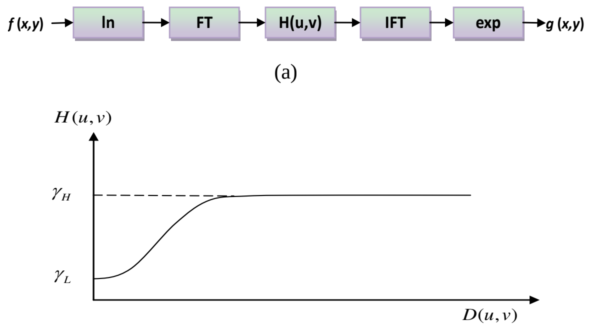
A. Homomorphic Filtering

Homomorphic filtering [22] is a classical tool used in image processing, attempting to normalize the brightness across an image and increase contrast via high-boost filtering in logarithm domain. The whole procedure of homomorphic filtering is illustrated in Fig. 2.

According to the Lambertian reflectance function, the intensity of a 2D surface I can be described as

$$I(x, y) = R(x, y)L(x, y) \tag{1}$$

where I, L, R represent the intensity, illumination component and reflectance component respectively. Note that, it is assumed that L changes very slowly [11], [17], [18], so the illumination effect mainly lies in the low-frequency



(b)

Fig. 1. Homomorphic filtering: (a) flowchart of the homomorphic filtering; (b) illustration of high-boost filter $H(u,v)$ used in (a).

domain. Thus, it is possible to alleviate the illumination effect by taking a high-pass or high-boost filtering in the frequency domain. However, as we can see in (1), it is not feasible to apply filtering on L since the reflectance component R and illumination component L are combined in the multiplicative form. Nevertheless, we find that the homomorphic filtering takes the logarithm transformation at the first step, making the two components combined in the additive form, so that the high-boost filtering on I can be separated into the sum of the high-boost filtering on L and that of R in the logarithm domain. Thus, applying homomorphic filtering will contribute to alleviate the effect caused by illumination variation:

$$\begin{aligned} &H(u,v)\mathscr{F}(\tilde{I}(u,v)) \\ =&H(u,v)\mathscr{F}(\tilde{F}(x,y)) + H(u,v)\mathscr{F}(\tilde{L}(x,y)) \end{aligned} \tag{2}$$

where $\tilde{I}(x,y)$, $\tilde{R}(x,y)$ and $\tilde{L}(x,y)$ represent $\ln(I(x,y))$, $\ln(R(x,y))$ and $\ln(L(x,y))$ respectively, $\mathscr{F}(\cdot)$ denotes the Fourier Transform (FT), and $H(u,v)$ is the filtering function.

Note that the high-frequency components are assumed to represent the reflectance mostly, whereas the illumination effect is mainly assumed to lie in the low-frequency domain [22]. Therefore, high-boost filtering can be applied here to suppress low frequency illumination effect and amplify high frequency facial characteristics. Specifically, we adopt the following kind of high-boost filter:

$$H(u,v) = (\gamma_H - \gamma_L)(1 - e^{-\frac{(u-u_0)^2+(v-v_0)^2}{\sigma^2}}) + \gamma_L, \tag{3}$$

where (u_0, v_0) represents the center location, γ_H, γ_L and σ are parameters to control the filter. After the filtering processing, we can obtain the enhanced images in the spatial domain via Inverse Fourier Transform (IFT). That is,

$$\hat{I}(x,y) = \mathscr{F}^{-1}(H(u,v)\mathscr{F}(\tilde{I}(u,v))) \tag{4}$$

The whole procedure of homomorphic filtering is illustrated in Fig. 2. Note that, different from the general homomorphic filtering which takes exponential transform on $\hat{I}(x,y)$ at the final step, we are going to extract the illumination invariant feature on $\hat{I}(x,y)$ directly.

B. Logarithm Gradient and its Illumination Invariant Property: Homogeneous Lighting

From observation, the shapes, the contours and the relative small-scale facial objects such as eyes, noses, mouths, etc, can be key features for face recognition [9]. The gradient information (e.g., magnitude and orientation) around these components contain much more valuable information than that of the skin areas. More importantly, the gradient magnitudes of these parts are always large and fluctuant while those of the skin areas are relative small. As a result, it is feasible to take the gradient magnitudes as importance measurement of facial components. However, as mentioned in Gradient Face [11], the gradient magnitudes of original face image do not satisfy the illumination invariant requirement. Indeed, this problem can be solved by transferring the derivation into logarithm domain so that the multiplicative combination in the original domain will become the additive form. Thus, different from Gradient Face, which only retain gradient orientation as the illumination invariant feature, we attempt to incorporate both gradient magnitude and direction in the logarithm domain to generate our illumination invariant features. The illumination invariant property of our proposed feature is based on the following theorems.

Lemma 1. *Assume $I(x,y)$ is an illuminated image of Lambertian face, which is captured under the homogeneous lighting (i.e., being with arbitrary lighting direction and magnitude but with the same lighting wavelength). Let $\tilde{I}(x,y) = ln(I(x,y))$, then the partial derivation $\partial_x \tilde{I}(x,y)$ and $\partial_y \tilde{I}(x,y)$ are insensitive to the illumination variation.*

Proof: For any function $f(x,y) > 0$, we denote $\tilde{f}(x,y) = ln(f(x,y))$. Let us consider two neighboring points (x,y) and $(x+\Delta x, y)$. According to the Lambertian Law, we have

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

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

then

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

$$\tilde{I}(x+\Delta x, y) = \tilde{R}(x+\Delta x, y) + \tilde{L}(x+\Delta x, y) \tag{8}$$

thus

$$\begin{aligned} &\tilde{I}(x+\Delta x, y) - \tilde{I}(x,y) \\ =&(\tilde{R}(x+\Delta x, y) - \tilde{R}(x,y)) + (\tilde{L}(x+\Delta x, y) - \tilde{L}(x,y)) \end{aligned} \tag{9}$$

As commonly assumed, L varies very slowly while R can change abruptly. The same assumption holds for $\tilde{L}$ and $\tilde{R}$. Therefore, as similarly suggested in [11], [17], [18], it is reasonable to draw the conclusion that the difference between $\tilde{L}(x+\Delta x, y)$ and $\tilde{L}(x,y)$ can be ignorable comparing to that between $\tilde{R}(x+\Delta x, y)$ and $\tilde{R}(x,y)$ when Δx is small enough. As a result, take the limit of both sides in Eq. (9), we have the following approximation:

$$\partial_x \tilde{I}(x,y) \approx \partial_x \tilde{R}(x,y) \tag{10}$$

Also, following the same procedure we can obtain

$$\partial_y \tilde{I}(x,y) \approx \partial_y \tilde{R}(x,y) \tag{11}$$

In a word, the partial derivation $\partial_x \tilde{I}(x,y)$ and $\partial_y \tilde{I}(x,y)$ are dominated by the reflectance component R instead of the illumination component L and thus insensitive to the illumination variation. ■

Theorem 1. *Assume $I(x,y)$ is an illuminated image of Lambertian face, which is captured under the homogeneous lighting (i.e., being with arbitrary lighting direction and magnitude but with the same lighting wavelength). Let $\tilde{I}(x,y) = ln(I(x,y))$, then the gradient orientation and gradient magnitude of $\tilde{I}(x,y)$ are both illumination invariant components.*

Proof: Using the same denotation in Lemma. 1, it is easy to calculate the gradient orientation and magnitude of $\tilde{I}(x,y)$ as

$$LGO(x,y) = \arctan(\partial_y \tilde{I}(x,y)/\partial_x \tilde{I}(x,y)) \tag{12}$$

and

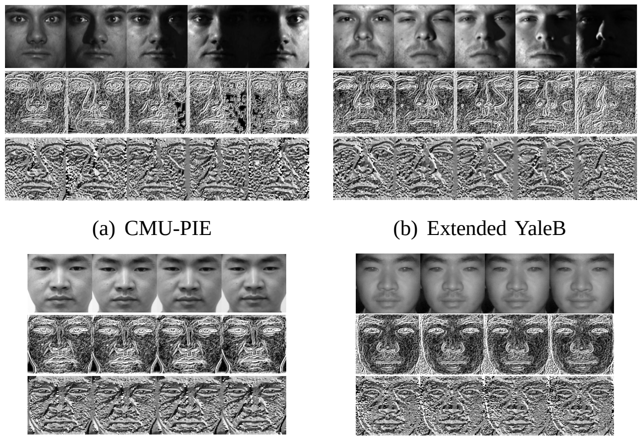
$$LGM(x,y) = \sqrt{(\partial_x \tilde{I}(x,y))^2 + (\partial_y \tilde{I}(x,y))^2} \tag{13}$$

As proved in Lemma. 1, $\partial_x \tilde{I}(x,y)$ and $\partial_y \tilde{I}(x,y)$ are both insensitive to illumination change, so that it is straightforward to get the conclusion that $LGO(x,y)$ and $LGM(x,y)$ are both illumination invariant components. ■

So far we have known that the the gradient orientation $LGO(x,y)$ and gradient magnitude $LGM(x,y)$ of $\tilde{I}(x,y)$ are both invariant to the illumination changes caused by varying lighting direction and magnitude. However, as previously mentioned, "the same lighting wavelength" assumption in homogeneous lighting can hardly hold in real world applications since the spectral wavelengths will change with environment. In the next section, we will further investigate the illumination invariant property of our proposed LGO and LGM when relaxing the assumption on spectral wavelength.

C. Illumination Invariant Property: Heterogeneous Lighting

According to the Lambertian model, the reflectance component, which is determined by the albedo (relate to spectral wavelength) and facial normal direction, will turn out to be different when suffering from heterogeneous lightings. Generally, it is not feasible to tackle this problem under wild assumption. For example, the facial skin reflectance under visible and far infrared light differ so much that the obtained face images are hardly the same. Nevertheless, according to the work on skin reflectance spectra simulation [23], we find that the responds of skin reflectance spectra change smoothly as the lighting wavelength increases within the visible and near-infrared spectral regions (450nm-1100nm). Inspired by this observation, we assume that within a small patch of facial skin, the reflectance components under two different lights (with different wavelengths) are proximately proportional. We denote it as the *locally proportional reflectance assumption* (see Theorem 2). Therefore, the key to solving such a general illumination problem is to develop an illumination invariant feature under this assumption. In the following part, we further prove that our proposed LGO and LGM are also invariant components in such scenario.



(a) CMU-PIE (b) Extended YaleB

(c) HFB VIS (d) HFB NIR

Fig. 2. Illustration of the logarithm gradient components. From top to bottom: original face images, the corresponding logarithm gradient magnitudes (LGM) and logarithm gradient orientations (LGO).

Theorem 2. *Assume $I_1(x,y) = R_1(x,y)L_1(x,y)$ and $I_2(x,y) = R_2(x,y)L_2(x,y)$ are Lambertian face images of the same subject captured under two heterogeneous lightings (with different lighting directions, magnitudes and even spectral wavelengths). Let $\tilde{I}(x,y) = ln(I(x,y))$, if $R_1(x,y)$ is locally proportional with $R_2(x,y)$, i.e., $R_1(x,y) = wR_2(x,y)$ for some constant w which is determined by $\mathcal{N}(x,y)$ (the neighborhood of (x,y)), then the gradient orientation and gradient magnitude of $\tilde{I}_1(x,y)$ and $\tilde{I}_2(x,y)$ are equal, i.e., they are both invariant components for the illumination problem.*

Proof: As indicated in Theorem. 1, the gradient orientation and gradient magnitude of $\tilde{I}(x,y)$, i.e., $LGO(x,y)$ and $LGM(x,y)$ are both invariant to the varying illumination direction and magnitude. Therefore, in this part we only need to prove the invariant property against the varying wavelength. By substituting (10) and (11) into (12) and (13), we have

$$\begin{aligned} LGO(x,y) &= \arctan(\frac{\partial_y \tilde{I}(x,y)}{\partial_x \tilde{I}(x,y)}) = \arctan(\frac{\partial_y \tilde{R}(x,y)}{\partial_x \tilde{R}(x,y)}) \\ &= \arctan(\frac{\partial_y R(x,y)/R(x,y)}{\partial_x R(x,y)/R(x,y)}) = \arctan(\frac{\partial_y R(x,y)}{\partial_x R(x,y)}) \end{aligned} \tag{14}$$

and

$$\begin{aligned} LGM(x,y) &= \sqrt{(\partial_x \tilde{I}(x,y))^2 + (\partial_y \tilde{I}(x,y))^2} \\ &= \sqrt{(\partial_x \tilde{R}(x,y))^2 + (\partial_y \tilde{R}(x,y))^2} \\ &= \frac{\sqrt{(\partial_x R(x,y))^2 + (\partial_y R(x,y))^2}}{R(x,y)} \end{aligned} \tag{15}$$

Assume that $R_1(x,y)$ and $R_2(x,y)$ are locally proportional.

That is,

$$R_1(x,y) = wR_2(x,y) \tag{16}$$

for some constant $w > 0$ which is determined by $\mathcal{N}(x,y)$. Thus,

$$\arctan(\frac{\partial_y R_1(x,y)}{\partial_x R_1(x,y)}) = \arctan(\frac{\partial_y R_2(x,y)}{\partial_x R_2(x,y)}) \tag{17}$$

i.e.,

$$LGO_1(x,y) = LGO_2(x,y) \tag{18}$$

Also, we have

$$\begin{aligned} LGM_1(x,y) =& \frac{\sqrt{(\partial_x R_1(x,y))^2 + (\partial_y R_1(x,y))^2}}{R_1(x,y)} \\ =& \frac{\sqrt{(k\partial_x R_2(x,y))^2 + (k\partial_y R_2(x,y))^2}}{kR_2(x,y)} \\ =& \frac{\sqrt{(\partial_x R_2(x,y))^2 + (\partial_y R_2(x,y))^2}}{R_2(x,y)} \\ =& LGM_2(x,y) \end{aligned} \tag{19}$$

Hence, it is guaranteed to draw the conclusion that $LGO(x,y)$ and $LGM(x,y)$ are both insensitive to the illumination variation caused by spectral wavelength change under the local proportional assumption. ■

D. The Logarithm Gradient Histogram

Until now, we have proposed a pair of illumination invariant components by full use of the gradient information in the logarithm domain, i.e., LGO and LGM, which are proved to be insensitive to the change of illumination direction, magnitude and even spectral wavelength within the specific regions. However, it is not reliable to directly conduct pixel-wise matchings on these two components among different face images. Hence, in the following part, we are going to integrate LGO and LGM into a unified histogram based feature representation.

To obtain a robust descriptor, postprocessing should be taken on the two components before generating the histogram representation. Consider that the gradient orientation is somehow sensitive to the quality of image, a local smoothing operation is suggested to be taken on LGO to alleviate the impulse responses caused by discrete noises and ensure the gradient direction changes smoothly. Taking interpolation or using a suitable σ for gaussian kernel as [11] are both feasible alternatives in this situation. After that, we quantify the values in LGO into several bins to achieve fault-toleration and prepare to generate gradient histogram. Meanwhile, it is worth mentioning that there are cast shadows in face images since the Lambertian assumption does not strictly hold everywhere. As a result, the pixels with dominating values in LGM may belong to the boundaries of shadows, and meanwhile the edges of facial objects (e.g., eyes, mouths) may become less significant especially when the lighting condition becomes severe. Thus, in this case, partial facial information part of information in LGM is not related to the and regarded as outlier. Thus taking the local normalizing operation here is able to restrain this effect.

Algorithm 1 Generating Logarithm Gradient Histogram

Input: LGO, LGM, k (the number of bins in histogram encoding each block)

Output: H (histogram of a face image)

1: Quantify LGO in k bins $\{b_1, \cdots, b_k\}$ to increase the fault-tolerant capability;

2: Divide LGM and LGO into small blocks evenly, and denote M^i and O^i as the ith block in LGM and LGO separately;

3: For each M^i, calculate the weighted normalized gradient magnitude as $\tilde{M}^i(p,q) = \frac{M^i(p,q)W(p,q)}{\sum_{(p',q')\in N(p,q)} M^i(p',q')W(p',q')}$, or else $\tilde{M}^i(p,q) = M^i(p,q)$, where W denotes the gaussian weight matrix.

4: After that, we can generate the histogram for each block, $H^i(t) = sum\{\tilde{M}^i(j)|O^i(j) == b_t\}$, for $t = 1, \cdots, k$, and then concatenate them into a long vector $H = [H^i]$ to represent a face image.

Finally, we obtain the quantified gradient orientation and normalized gradient magnitude for each pixel in a face images. The histogram generating procedure is implemented in a block-wise form, that is, the post-processed gradient magnitudes of all pixels in the block will accumulate according to the orientation bins they belong. At last, we concatenate histograms of all blocks into a long vector and form our histogram based feature representation. The whole procedure is illustrated in Algorithm 1.

IV. Experiments

In this section, we conducted a series of experiments to evaluate the proposed illumination invariant descriptor. Two scenarios will be considered in our experiment: 1) the case when images were approximately captured under the homogeneous lighting, i.e., with different lighting directions and magnitudes but the same spectral wavelength; 2) the case when image were captured under heterogeneous lighting, i.e., with varying lighting direction, magnitude and wavelength. In literature, different methods have developed to solve different case. Hence, for the first case, we compare our method with HE [3], SQI [8], LTV [9], Weber Face [12] and Gradient Face [11], on both CMU-PIE and Extended YaleB databases; for the second case, we particular consider the heterogeneous face recognition and compare our method with DoG+LBP [15] and SIFT descriptor used in [16] on the heterogeneous face biometric (HFB) database [24]. In our experiment, we mainly evaluate the performances of different image descriptors and thus simply adopted the generic Euclidean distance as the similarity measurement and used nearest neighborhood classifier for classification.

A. Face Recognition: Homogeneous Lighting

1) Databases and Setting: In this experiment, we evaluated the performance of various methods on the CMU-PIE and the Extended YaleB databases. To form the set of frontal face images, 1428 frontal face images from 68

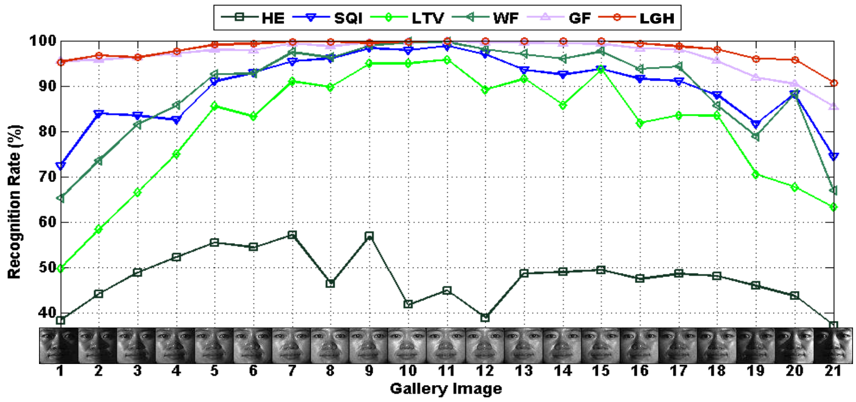


Fig. 3. Recognition rates with different galleries on CMU-PIE

individuals under 21 different illumination conditions were selected from the CMU-PIE database. For Extended YaleB, face images from 38 individuals were captured under 64 different lighting conditions on 9 poses, and we only used $64 \times 38 = 2432$ frontal face images here. All images are simply aligned according to the eyes coordinates and resized to 128×128.

we compared our LGH to several classical and state-of-the-art methods, including HE [3], SQI [8], LTV [9], Weber Face [12] and Gradient Face [11] on both CMU-PIE and Extended YaleB databases. All methods were implemented with parameters set as suggested in the references. For our LGH, the parameters of homomorphic filtering are empirically fixed as $\gamma_H = 2.0$, $\gamma_L = 0.5$ and $\sigma = c * L$, where L is width of the image, c is a constant to control the radius of high-pass filter and fixed as 0.1 in all experiments. Unless otherwise stated, the block size is set as 4×4, unsigned gradient orientation is adopted and 5 bins are used for the quantization procedure.

2) Results on CMU-PIE: The subset used here consists of 21 images for each person, which are captured under 21 different illumination conditions as shown in Fig. (3). Only one image per individual was chosen as gallery and the other formed the probes. We varied the gallery from the 1st image to the 21st one for each person in order to ensure that all illumination conditions were covered. The final results are illustrated in Fig. 3. Note that, we have resorted the images to make the light source change from lhs to rhs gradually. As we can see, the general performances of different algorithms degenerated as the lighting orientation diverges from the frontal direction. All six methods except HE achieve the best performances when using the frontal lighting images as galleries, and the corresponding results for various methods are as follows: 98.82% for SQI, 95.81% for LTV, 99.71% for WF, 99.93% for GF and 100% for LGH. What's more, some approaches such as SQI, LTV and WF turn out less effectively when the illumination condition become severe, while both GF and LGH perform well even under the most extremely situations. Also, our proposed LGH outperforms all other methods almost in each lighting condition and achieves the best average recognition rate at 98.19%. The average results for HE, SQI, LTV, WF and GF are 47.59%, 89.77%, 80.78%, 89.52% and 96.93% respectively.

TABLE I
RESULTS OF EXTENDED YALEB FOLLOWING PROTOCOL I (IN ACCURACY (%))

	Set1	Set2	Set3	Set4	Set5	All
HE [3]	**97.81**	92.76	36.18	10.90	13.43	41.25
SQI [8]	88.60	**100.00**	85.75	87.97	81.02	87.82
LTV [9]	87.28	99.78	66.67	45.49	44.32	63.86
WF [12]	79.39	99.78	75.88	77.07	74.38	80.56
GF [11]	94.74	**100.00**	83.33	75.94	74.65	83.51
LGH	94.74	**100.00**	**92.54**	**96.43**	**86.70**	**93.30**

TABLE II
RESULTS OF EXTENDED YALEB FOLLOWING PROTOCOL II (IN ACCURACY (%))

	Set1	Set2	Set3	Set4	Set5
HE [3]	66.61±31.37	60.26±21.49	53.34±10.98	45.37±12.72	54.29±14.78
SQI [8]	81.05±19.49	80.14±19.81	85.80±12.17	88.51±5.52	94.51±1.82
LTV [9]	79.85±22.92	78.47±22.89	70.31±16.43	56.57±8.62	73.34±8.16
WF [12]	90.38±8.17	86.68±13.13	90.01±8.65	88.58±3.82	93.61±3.21
GF [11]	94.32±7.73	91.40±10.78	91.21±7.68	89.86±4.34	96.13±5.82
LGH	**98.88±1.25**	**97.96±2.55**	**99.49±0.78**	**97.59±1.67**	**96.39±2.22**

3) Results on Extended YaleB: Different from CMU-PIE, the face images in Extended YaleB database were captured in more complex environments. To better explore the performance of our proposed illumination invariant descriptor, we conducted experiments following two protocols and reported recognition accuracies on set 1 to 5 separately.

Protocol I: the frontal lighting image per subject was chosen as gallery set and the rest for probe set;

Protocol II: three images under arbitrary lighting conditions were randomly chosen to form the gallery set and the rest formed the probe set.

It is worth mentioning that, the angles between frontal face directions and lighting orientations increase from set 1 to set 5. That is, generally, it is more challenging to handle the face recognition task as the set index increases, which

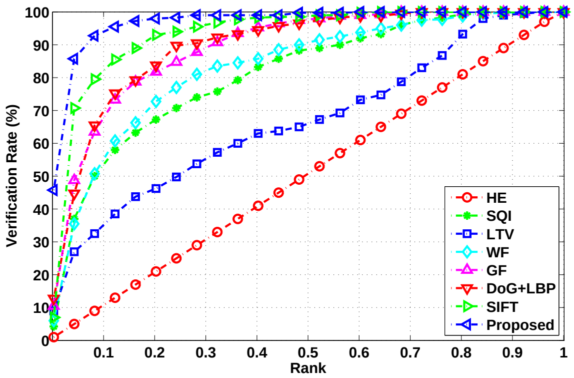


Fig. 4. Cumulative Match Curves (CMC) of the compared methods on the HFB VIS-NIR dataset.

will be reflected by results in Table I and Table II. Note that, results of *All* in Table I are calculated by taking all other 63 images per subject except the gallery one as probe. As shown in Table I, it is interesting to find that HE obtains the highest accuracy 97.81% in Set 1 where single frontal illuminated images was used as gallery, while GF and LGH achieve the same recognition rate at 94.74% only. It makes sense because the illumination effect is limited and the lighting is uniformly distributed all over the face in Set 1, so that the HE performs well in this scenario. For our proposed LGH, as shown in Table I, it outperforms the other methods in all sets except the case in Set 1 following protocol I.

However, the frontal lighting condition is sometimes relative strict for realistic applications. As a result, the gallery set may contain images under different lighting conditions. It is interesting to evaluate the robustness of illumination invariant descriptors in this scenario. Hence, we conducted another experiment using three randomly chosen images instead of the one with frontal illumination to consist the gallery set in protocol II. Both average recognition rates and corresponding standard deviations of 20 random trials are reported in this part, and all results are shown in Table II. We observe that our proposed LGH achieves consistent highest performance in this case where the average accuracies of set 1 to set 5 are 98.88%, 97.96%, 99.49%, 97.59% and 96.39% respectively, obtaining 4.83%, 7.18%, 9.08%, 8.60% and 0.27% improvement compared to the second best approach (i.e. CG).

B. Face Recognition: Heterogeneous Lighting

1) Databases and Setting: To further investigate the effectiveness of our proposed LGH in handling the more general case in illumination problem, we used the VIS-NIR subset of HFB database, which contains 100 people and each is with 4 VIS and 4 NIR face images. Note that, the spectral wavelength of NIR light used here is around 850nm while that of VIS light is smaller than 700nm. It is more challenging since they are captured under two distinct heterogeneous light sources with no overlapping wavelength. Meanwhile, there are still some variations caused by the facial expression and occultation. All images are aligned according to the eyes coordinates and resized to 128×128. For comparison, results of DoG+LBP [15] and SIFT descriptor used in [16] accompanied with methods mentioned in Section IV-A are also reported.

2) Results on HFB: In this part, the gallery set consists of all four VIS face images per individual while the probe one contains all other NIR face images. It is worth pointing out that no learning method is further used here. Two descriptors which are applied in VIS-NIR face matching were tested here for complemental comparison, denoting as DoG+LBP [15] and SIFT [16]. We report the final results in terms of cumulative match curve (CMC) as shown in Fig. (4).

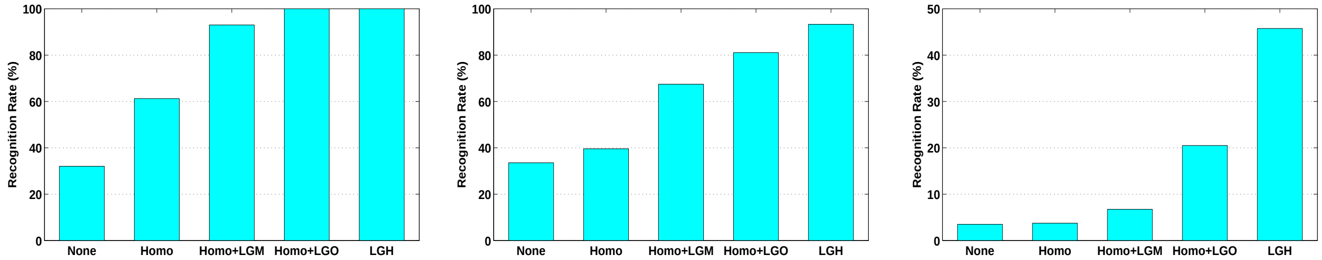
Note that, variations including those caused by expression and occultation (by glasses) also exist in this database, leading to a more challenging task beyond the illumination problem. As a result, the overall performances of all methods degenerate dramatically. The previously mentioned approaches as well as DoG+LBP and SIFT obtain low Rank-1 recognition rates at less than 15% while the result of our proposed method is up to 45.75% with a large improvement. Different from the case in existing approaches, the assumption of the same lighting wavelength no longer holds here, making the aforementioned descriptors (such as SQI, LTV, WF, etc) invalid. However, our proposed LGH, outperforms other methods and shows great tolerability to the general illumination change, which again confirms our previous analysis in Section III. In addition, as illustrated in Fig. (4), our LGH also achieves significant improvements even compared to the two empirically designed descriptors, i.e., DoG+LBP and SIFT.

C. Contribution of Each Component in LGH

Since LGH consists of three parts including homomorphic filtering (denoted as Homo), logarithm gradient magnitude (LGM) and logarithm gradient orientation (LGO), we would like to explore the contribution of each part. In this section, we conducted experiments on CMU-PIE, Extended YaleB and HFB VIS-NIR databases and reported the recognition rates of each part separately in Fig. (5). Note that, only one frontal illuminated face image was used as gallery in CMU-PIE and Extended YaleB, and all VIS images in HFB were chosen as gallery. The final results validate our previous analysis that (i) the homomorphic filtering is able to restrain the illumination effect in some respect; (ii) gradient magnitude and gradient orientation in logarithm are somehow insensitive to the illumination change; (iii) our proposed LGH integrates the above two components in an effective way and achieves great success in tolerating lighting changes, especially when the change is severe.

V. Conclusions

In this paper, we have proposed a novel illumination invariant descriptor LGH to address the illumination problem. Different from the existing methods, we consider variations caused by the lighting direction, magnitude and even the spectral wavelength. On the basis of illumination invariant property analysis, we develop two illumination invariant



(a) CMU-PIE (b) Extended YALEB (c) HFB

Fig. 5. Contribution of each part in LGH. Only one frontal illuminated face image was used as gallery in (a) and (b), all VIS images in HFB were chosen as gallery in (c). Notations used in x-axis from left to right represented original image (None), image after homomorphic filtering (Homo), logarithm gradient magnitude followed by homomorphic filtering (Homo+LGM), gradient orientation followed by homomorphic filtering (Homo+LGO) and the whole LGH (LGH).

components in the logarithm domain after homomorphic filtering, i.e., LGO and LGM. After that, we integrate them into a histogram based feature representation followed by post-processing to enhance the fault-tolerant ability. Experimental results verify the effectiveness of our proposed method in tackling with illumination problems under the setting from the homogeneous lighting so to heterogeneous lighting.

VI. Acknowledgments

This research was partially supported by Guangdong Provincial Government of China through the Computational Science Innovative Research Team program, the National Natural Science of Foundation of China (Nos.61102111, 61173084), Specialized Re-search Fund for the Doctoral Program of Higher Education (No.20110171120051) and the 985 Project at Sun Yat-sen University under Grant No.35000-3281305.

References

[1] P. Phillips, W. Scruggs, A. O'Toole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe, "Frvt 2006 and ice 2006 large-scale experimental results," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 5, pp. 831–846, 2010.

[2] X. Xie, J. Lai, and W. Zheng, "Extraction of illumination invariant facial features from a single image using nonsubsampled contourlet transform," *Pattern Recognition*, vol. 43, no. 12, pp. 4177–4189, 2010.

[3] S. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[4] S. Shan, W. Gao, B. Cao, and D. Zhao, "Illumination normalization for robust face recognition against varying lighting conditions," in *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*. IEEE, 2003, pp. 157–164.

[5] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 6, pp. 643–660, 2001.

[6] H. Shim, J. Luo, and T. Chen, "A subspace model-based approach to face relighting under unknown lighting and poses," *Image Processing, IEEE Transactions on*, vol. 17, no. 8, pp. 1331–1341, 2008.

[7] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 2037–2041, 2006.

[8] H. Wang, S. Li, and Y. Wang, "Generalized quotient image," in *Proceedings of the 2004 IEEE computer society conference on Computer vision and pattern recognition*. IEEE Computer Society, 2004, pp. 498–505.

[9] T. Chen, W. Yin, X. Zhou, D. Comaniciu, and T. Huang, "Total variation models for variable lighting face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 9, pp. 1519–1524, 2006.

[10] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Proceedings of the 3rd international conference on Analysis and modeling of faces and gestures*. Springer-Verlag, 2007, pp. 168–182.

[11] T. Zhang, Y. Tang, B. Fang, Z. Shang, and X. Liu, "Face recognition under varying illumination using gradientfaces," *Image Processing, IEEE Transactions on*, vol. 18, no. 11, pp. 2599–2606, 2009.

[12] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on weber's law with application to face recognition," *Signal Processing Letters, IEEE*, no. 99, pp. 1–1, 2011.

[13] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 2, pp. 218–233, 2003.

[14] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 1. Ieee, 2005, pp. 786–791.

[15] S. Liao, D. Yi, Z. Lei, R. Qin, and S. Li, "Heterogeneous face recognition from local structures of normalized appearance," *Advances in Biometrics*, pp. 209–218, 2009.

[16] B. Klare and A. Jain, "Heterogeneous face recognition: Matching nir to visible light images," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 1513–1516.

[17] D. Jobson, Z. Rahman, and G. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *Image Processing, IEEE Transactions on*, vol. 6, no. 7, pp. 965–976, 1997.

[18] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2003, pp. 1055–1055.

[19] Z. Hou and W. Yau, "Relative gradients for image lighting correction," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1374–1377.

[20] A. Jain, *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.

[21] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International journal of computer vision*, vol. 60, no. 1, pp. 63–86, 2004.

[22] M. SONKA, V. HLAVAC, and R. BOYLE, "Image processing, analysis and machine vision (1998)," *Recherche*, vol. 67, p. 02, 1998.

[23] I. Meglinski and S. Matcher, "Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions," *Physiological measurement*, vol. 23, p. 741, 2002.

[24] S. Li, Z. Lei, and M. Ao, "The hfb face database for heterogeneous face biometrics research," in *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*. IEEE, 2009, pp. 1–8.