Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Pattern Recognition Letters 27 (2006) 609–617
Pattern Recognition Letters
www.elsevier.com/locate/patrec

An efficient illumination normalization method for face recognition

Xudong Xie, Kin-Man Lam *

Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Received 11 December 2004; received in revised form 20 September 2005
Available online 15 November 2005

Communicated by Prof. G. Sanniti di Baja

Abstract

In this paper, an efficient representation method insensitive to varying illumination is proposed for human face recognition. Theoretical analysis based on the human face model and the illumination model shows that the effects of varying lighting on a human face image can be modeled by a sequence of multiplicative and additive noises. Instead of computing these noises, which is very difficult for real applications, we aim to reduce or even remove their effect. In our method, a local normalization technique is applied to an image, which can effectively and efficiently eliminate the effect of uneven illuminations while keeping the local statistical properties of the processed image the same as in the corresponding image under normal lighting condition. After processing, the image under varying illumination will have similar pixel values to the corresponding image that is under normal lighting condition. Then, the processed images are used for face recognition. The proposed algorithm has been evaluated based on the Yale database, the AR database, the PIE database, the YaleB database and the combined database by using different face recognition methods such as PCA, ICA and Gabor wavelets. Consistent and promising results were obtained, which show that our method can effectively eliminate the effect of uneven illumination and greatly improve the recognition results.

Keywords: Face recognition; Illumination compensation; Local normalization method; Principal component analysis (PCA); Independent component analysis (ICA); Gabor wavelets

1. Introduction

Human face recognition, as one of the most successful applications of image analysis and understanding, has received significant attention in the last decade (Zhao et al., 2003). However, given that, as mentioned in (Adini et al., 1997), "the variations between the images of the same face due to illumination and viewing direction are almost always larger than the image variations due to a change in face identity", most existing methods for face recognition, such as principal component analysis (PCA) (Sirovich and Kirby, 1987; Kirby and Sirovich, 1990; Turk and Pentland, 1991; Pentland, 2000) and independent component analysis (ICA) (Bartlett et al., 2002; Liu and Wechsler, 2003; Déniz et al., 2003; Draper et al., 2003), encounter difficulties under varying lighting conditions. Hence, if only one upright frontal image per person, which is under even illumination, is available for training, the performance of PCA and ICA will be seriously degraded if the testing faces are under severe lighting variations. In this paper, we will address the impact of varying illuminations on face recognition.

Many methods have been proposed to handle the illumination problem. The linear subspace method (Hallinan, 1994; Bichsel, 1995; Belhumeur et al., 1997) considered a human face image as a Lambertian surface, which can use three or more images of an object under different lighting conditions to compute a basis for the 3D illumination subspace. Without ignoring the shadows, the 3D illumination

* Corresponding author. Tel.: +852 2766 6207; fax: +852 2362 8439.
E-mail address: enkmlam@polyu.edu.hk (K.-M. Lam).

0167-8655/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.09.026

subspace model was extended to a more elaborate one, namely the illumination convex cone (Belhumeur and Kriegman, 1996; Georghiades et al., 1998, 2000). Ishiyama and Sakamoto (2002) proposed a geodesic illumination basis model, which calculates pose-independent illumination bases for a 3D model. Batur and Hayes (2001) presented a segmented linear subspace model by segmenting the images into regions that have surface normals with directions close to each other. Zhao and Yang (1999) attempted to account for the arbitrary effects of illumination on PCA-based vision systems by first generating an analytically closed-form formula of the covariance matrix of faces under a particular lighting condition, and then converting it to an arbitrary illumination via an illumination equation. All the above-mentioned methods usually require a set of known face images under different lighting conditions for training.

Zhao and Chellappa (2000) developed a shape-based face recognition system by means of an illumination-independent ratio image derived by applying a symmetrical shape-from-shading technique to face images. Chen et al. (2000) adopted a probabilistic approach in which a probability distribution for the image gradient is analytically determined. Shashua and Riklin-Raviv (2001) used quotient images to solve the problem of class-based recognition and image synthesis under varying illumination. Zhao et al. (2003) proposed illumination ratio images, which can be used to generate new training images for face recognition with a single frontal view image. Xie and Lam (2005a) proposed a model-based illumination compensation scheme for face recognition, which adopts a 2D face shape model to eliminate the effect of difference in the face shape of different persons. Liu et al. (2005) also proposed a method that can restore a face image captured under an arbitrary lighting condition to the one with frontal illumination by using a ratio image.

In this paper, a novel illumination normalization method for human face recognition is proposed. In our method, a human face is treated as a combination of a sequence of small and flat facets. The effect of the illumination on each facet is modeled by a multiplicative noise and an additive noise. Therefore, a local normalization (LN) technique (Xie and Lam, 2005b) is applied to the image, which can effectively and efficiently eliminate the effect of uneven illumination. Then the generated images, which are insensitive to illumination variations, are used for face recognition using different methods, such as PCA, ICA and Gabor wavelets (Chui, 1992; Liu et al., 2004).

This paper is organized as follows. In Section 2, the human face and illumination models adopted in this paper are introduced. The LN method, which is used to eliminate the effect of uneven illuminations, is presented in Section 3. In Section 4, experimental results are detailed and the use of different illumination compensation/normalization algorithms with different face recognition algorithms based on different databases are evaluated. Finally, in Section 5, conclusions are drawn.

2. Human face model and illumination model

A face image is supposed to be a Lambertian surface, which can be described as the product of the albedo and the cosine angle between the point light source and the surface normal as follows:

$$I(x,y) = \rho(x,y)\boldsymbol{n}(x,y)\cdot\boldsymbol{s}, \tag{1}$$

where $I(x,y)$ is the intensity value of the pixel at (x,y) in the image, $0 \leqslant \rho(x,y) \leqslant 1$ is the corresponding albedo, $\boldsymbol{n}(x,y)$ is the surface normal direction, $\boldsymbol{s}$ is the light source direction, and its magnitude is the light source intensity.

In computer graphics applications, a human face is treated as a combination of a sequence of small and flat facets (Feng and Yuen, 2000; Hwang and Lee, 2003), which can be determined by important facial feature points. Fig. 1 shows a face image overlaid with an updated version of the CANDIDE model (Ahlberg, 2001), which is composed of a sequence of triangular facets.

The area of each facet W is small enough to be considered a planar patch. Therefore, for each point $(x,y)\in W$, the surface normal direction $\boldsymbol{n}(x,y)$ is a constant. Furthermore, we assume that the light source used is directional, and therefore a good approximation of real situations (Zhao and Yang, 1999). Thus, the light source direction $\boldsymbol{s}$ is almost constant within W. Then, from (1), it is clear that the intensity value of the pixel at (x,y) is equal to the multiplication of the albedo at (x,y) and a scalar, which is constant within W. Suppose $f(x,y)$ and $f'(x,y)$ represent the pixel intensity values at (x,y) of the image under normal lighting conditions and the image under a certain kind of illumination, and $\boldsymbol{s}$ and $\boldsymbol{s}'$ are the corresponding light source directions. Then the corresponding illumination ratio image (Zhao et al., 2003) is given as follows:

$$\begin{aligned} R_i &= f'(x,y)/f(x,y) \\ &= (\rho(x,y)\boldsymbol{n}(x,y)\cdot\boldsymbol{s}')/(\rho(x,y)\boldsymbol{n}(x,y)\cdot\boldsymbol{s}) \\ &= (\boldsymbol{n}(x,y)\cdot\boldsymbol{s}')/(\boldsymbol{n}(x,y)\cdot\boldsymbol{s}) = A, \quad (x,y)\in W, \end{aligned} \tag{2}$$

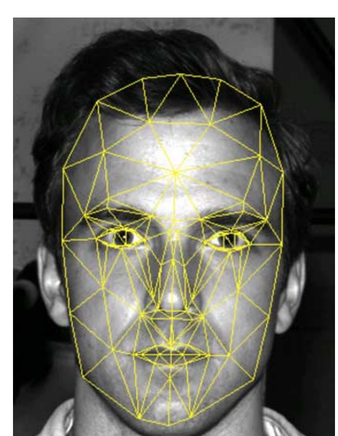

Fig. 1. A human face image and its corresponding CANDIDE-3 model.

where A is determined by the surface normal direction $\boldsymbol{n}$ of W and the kind of illumination concerned. For a special kind of illumination, the value of A is fixed within the facet W. From (2), we can obtain:

$$f'(x,y) = A \cdot f(x,y), \quad (x,y) \in W. \tag{3}$$

If we consider the effect of noise at each point $(x,y) \in W$, the illumination model in (3) can be extended to the following:

$$f'(x,y) = A \cdot f(x,y) + B, \quad (x,y) \in W, \tag{4}$$

where A and B denote the multiplicative noise and the additive noise for the pixel (x,y), respectively, and they are constant within W. In (4), $f'(x,y)$ is the intensity value at (x,y). A and B are unknown, and the problem is how, given $f'(x,y)$, to estimate the intensity value $f(x,y)$ of the face image under normal illumination. This is an ill-posed problem. Although we assume that the values of A and B are constant in a facet W, the real range of W is unknown as it depends on the shape of a face image and is difficult to obtain under varying illumination. In (Xie and Lam, 2005a), a 2D face shape model was adopted to map an image into a shape-free texture, and the YaleB (Yale University, 2001) database was then used to form the training set to obtain the A and B values pixel by pixel for each lighting category. In this paper, instead of estimating the values of A and B, we eliminate the effect of A and B by using the local normalization technique.

3. Local normalization technique

The main idea behind the LN technique is that, after processing an image $f'(x,y)$, its intensity value $f'_P(x,y)$ is of local zero mean and with unit variance within a facet W, i.e.

$$E(f'_P(x,y)) = 0 \quad \text{and} \quad \mathrm{Var}(f'_P(x,y)) = 1,$$
$$\text{where} \quad (x,y) \in W. \tag{5}$$

We define

$$f'_P(x,y) = \frac{f'(x,y) - E(f'(x,y))}{\mathrm{Var}(f'(x,y))}, \quad (x,y) \in W, \tag{6}$$

where $E(f'(x,y))$ is the mean of $f'(x,y)$ within W and $\mathrm{Var}(f'(x,y))$ is the corresponding variance. Then, from (4), we have

$$\begin{aligned} E(f'(x,y)) &= E(A \cdot f(x,y) + B) \\ &= A \cdot E(f(x,y)) + B, \quad (x,y) \in W \end{aligned} \tag{7}$$

and

$$\begin{aligned} \mathrm{Var}(f'(x,y)) &= \sqrt{\frac{\sum (f'(x,y) - E(f'(x,y)))^2}{N}} \\ &= A \cdot \sqrt{\frac{\sum (f(x,y) - E(f(x,y)))^2}{N}} \\ &= A \cdot \mathrm{Var}(f(x,y)), \quad (x,y) \in W, \end{aligned} \tag{8}$$

where N is the number of pixels within W, $E(f(x,y))$ and $\mathrm{Var}(f(x,y))$ are the corresponding local mean and local variance of $f(x,y)$. From 4 and (6)–(8), we have

$$f'_P(x,y) = \frac{f(x,y) - E(f(x,y))}{\mathrm{Var}(f(x,y))}, \quad (x,y) \in W. \tag{9}$$

In order to avoid overflow, a small constant (equal to 0.01) is added to all the variance values, which does not affect the derivation of (9). The image $f'_P(x,y)$ satisfies the conditions in (5), as proved in Appendix A. Furthermore, it is obvious that after the LN processing, the intensity value of the pixel at (x,y) is determined only by the corresponding intensity value $f(x,y)$ of the image, which is under normal illumination, and the local statistical properties of $f(x,y)$. In other words, the effects of the uneven illumination, namely the multiplicative noise A and the additive noise B, can be eliminated completely.

As with an image $f(x,y)$ under normal illumination, after the local normalization, we have

$$f_P(x,y) = \frac{f(x,y) - E(f(x,y))}{\mathrm{Var}(f(x,y))}, \quad (x,y) \in W. \tag{10}$$

From (9) and (10), we can obtain that

$$f'_P(x,y) = f_P(x,y), \quad (x,y) \in W. \tag{11}$$

This means that, after the LN processing, the image under varying illumination will have the same intensity values as the image under normal lighting conditions. This property is very useful, and we can use the images, after LN processing, for face recognition.

Our discussion in this paper is based on the assumption that a human face can be considered a combination of a sequence of small and flat facets. Within each facet, applying the LN technique can obtain the illumination insensitive property for each pixel. However, it is difficult to determine the range or size of a facet, especially for images under varying illuminations. In our method, we simply apply a filter of size $N \times N$ to each pixel. In other words, the filter is centered on the pixel under consideration and the corresponding mean and variance of the pixel intensities within the window are computed, then (6) is applied to normalize the intensity of the pixel. This process is repeated pixel by pixel to obtain a representation that is insensitive to lighting.

In (6), the local mean and variance of an image are computed point by point. The images formed by the local means and variances, denoted as $E(f(x,y))$ and $\mathrm{Var}(f(x,y))$, are called the local mean and variance maps, respectively. Fig. 2 illustrates some original images in the YaleB database in the first row, those images processed by histogram equalization (HE) in the second row, the corresponding local mean maps and local variance maps in the third and fourth rows, respectively, and those processed by our LN algorithm in the last row. For Fig. 2(c)–(e), the block size used is 7×7 for local normalization.

Fig. 2 shows that the local mean map of an image represents its low-frequency contents, while the local variance

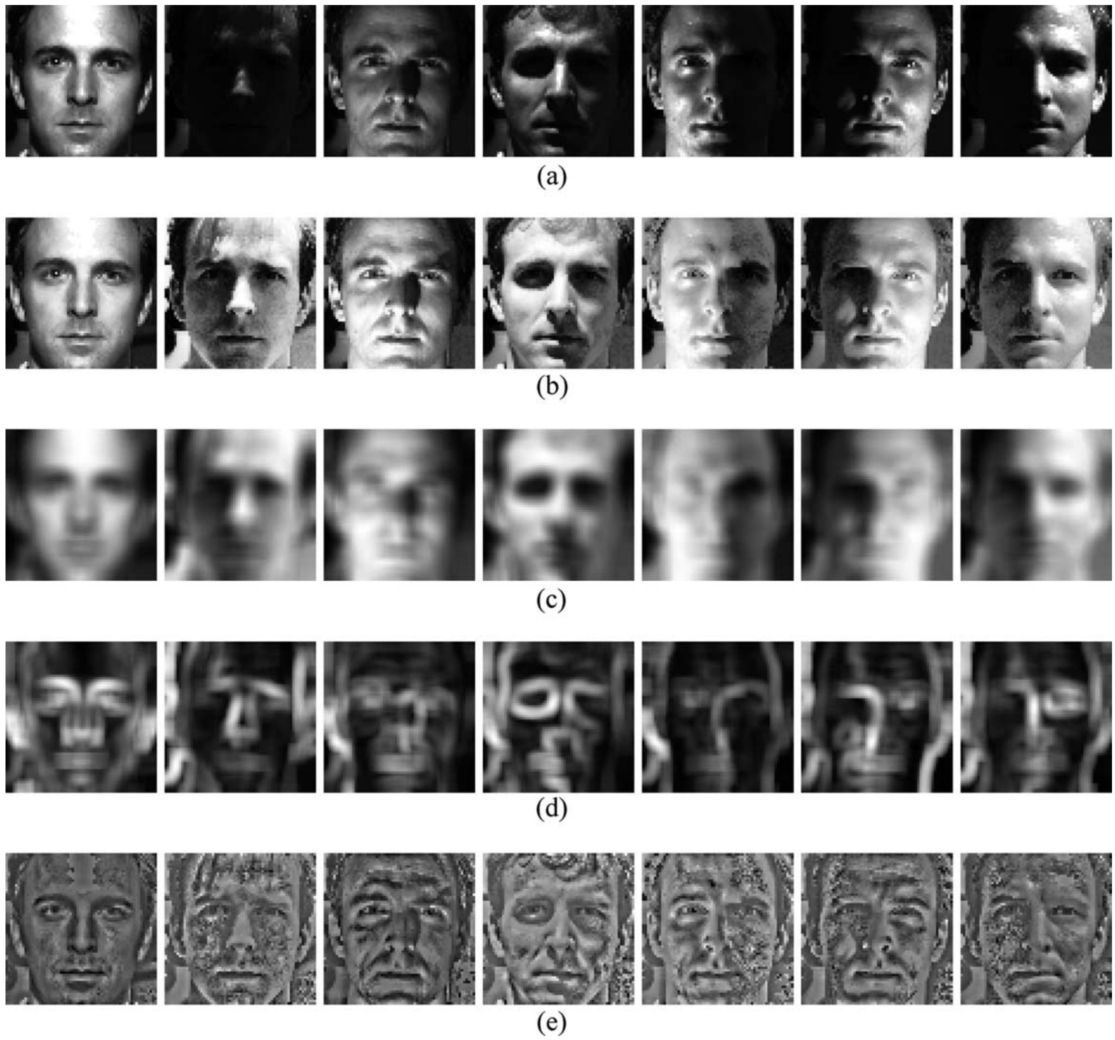

Fig. 2. Samples of cropped faces used in our experiments. The azimuth angles of the lighting of images from left to right column are: 0°, 0°, 20°, 35°, 70°, −50° and −70°, respectively. The corresponding elevation angles are: 20°, 90°, −40°, 65°, −35°, −40° and 45°, respectively. (a) Original images, (b) images processed using histogram equalization, (c) local mean maps, (d) local variance maps, (e) images processed using LN.

map carries the high-frequency components, or more accurately, the edge information about the image. This is because those pixels that lie in edge areas should have higher local variance values, and vice versa. In the case of uneven lighting conditions, the local mean maps are dominated by the varying illuminations, and the edge information is disturbed by the varying local contrast and shadows. Therefore, from (6), we can see that in the local normalization process, the subtraction of an image by its local mean map can reduce the global uneven lighting effect, and then dividing it by its local variance map can further reduce the effect of unreliable edge information. In other words, after these two procedures, the effects of uneven illumination on both the low-frequency and high-frequency components of an image will be reduced or even eliminated. The processed image becomes robust to illumination variation and can therefore be used to achieve a more reliable performance for face recognition.

4. Experimental results

In this section, we will evaluate the performance of the LN algorithm for face recognition based on different face databases. The databases used include the Yale database (Yale University, 1997), the AR database (Martinez and Benavente, 1998), the YaleB database and the PIE database (Sim et al., 2002). We have also combined the four databases in the experiments. The number of distinct subjects and the total number of testing images in the respective databases are tabulated in Table 1.

Table 1
The subjects and test set images in the experiments

	Yale	AR	YaleB	PIE	Combined
Subject	15	121	10	68	214
Testing set	30	363	640	1564	2597

For each database, the lighting conditions are different. In the Yale database, the lighting is either from the left or the right of the face images. In the AR database, besides the lighting from the left and the right, there is also lighting from both sides of a face. The YaleB database, which consists of 10 people with 65 images of each person under different lighting conditions, is often used to investigate the effect of lighting on face recognition. In the PIE database, 24 different illumination models are adopted.

All images are cropped and normalized to a size of 64×64, and are aligned based on the two eyes. In our system, the position of the two eyes can be located either manually or automatically (Wong et al., 2001; Lam and Yan, 1996), and the input color images are converted to gray-scale ones. In the experiments, the eyes are located manually. If the eyes are detected automatically, the recognition rates of the respective methods considered will degrade due to the error in detecting the eyes. Our method is based on the local statistical properties of images. Therefore, in order to reduce the effect of pepper noise, a 3×3 filter is adopted to detect any isolated noise point, whose intensity value will then be replaced by the mean value of the pixels within its 3×3 neighborhood.

4.1. The block size for local normalization

The block size used in the LN process will affect the performance in compensating for the illumination effect and, thus, the rate for face recognition. Fig. 3 shows some images processed using the LN method with different block sizes. When the block size is very small, the statistical parameters $E(f(x,y))$ and $\mathrm{Var}(f(x,y))$ at (x,y) are not reliable, and the output images will be noisy. However, if the block size is too large, the assumption that all the pixels within a block are located within a facet is no longer tenable, and the illumination insensitive property of the processed images also becomes invalid. Therefore, an appropriate block size is important for LN processing.

In order to select a proper block size, PCA is used for face recognition with images processed using the LN method with different block sizes. In order to enhance the global contrast on the input images, histogram equalization is also adopted for image preprocessing (Section 4.2 will provide a more detailed discussion of the effect of histogram equalization). In other words, all images are first processed by histogram equalization and local normalization sequentially, and are then followed by feature extraction and face recognition using the PCA method. Fig. 4 shows the recognition rates based on different databases. For each database, with an increase of the block size, the recognition rate will rapidly increase until the block size reaches a critical value. Then, the recognition rate will decrease slowly. The critical or optimal filter size varies for different databases; each database has distinct characteristics in terms of the lighting conditions. We can see that the Yale and AR databases are more sensitive to the block size compared to the other databases, and the PIE database is almost independent of the window size. In our algorithm, we set the block size at 7×7, at which the combined database can obtain the best recognition rate.

4.2. Face recognition based on different databases

In this section, we will evaluate the performances of different lighting compensation/normalization methods for different face recognition techniques such as PCA, ICA and Gabor wavelets. The lighting compensation/normalization schemes evaluated in the experiments include the histogram equalization (HE) method, our proposed local normalization (LN) method, and the use of both HE and LN, i.e. HE + LN. We use the databases shown in Table 1 for testing. In each database, one frontal image of each subject with normal illumination and neutral expression was selected as a training sample, and others form the testing set.

4.2.1. Face recognition using PCA

PCA is a classical method of human face representation and recognition. The major idea of PCA is to represent faces with a weighted sum of a small collection of the principal components of the training images, namely

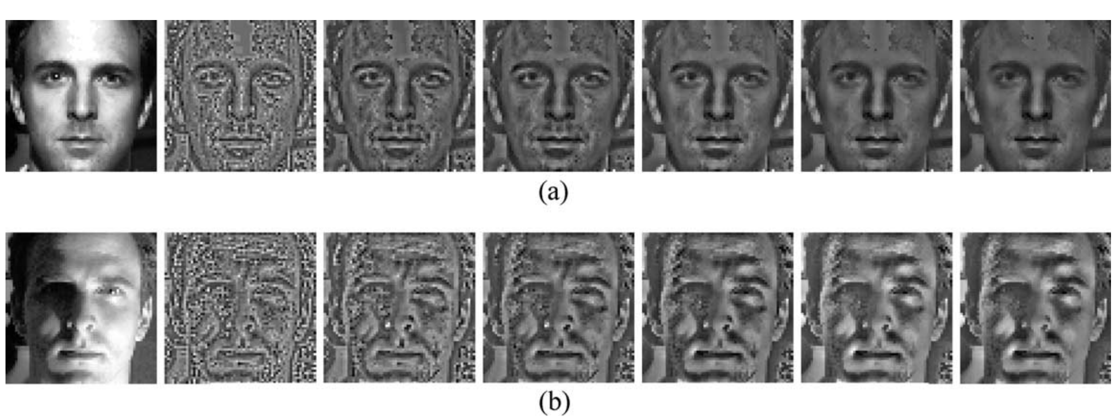

Fig. 3. Face images processed using the LN technique with different block sizes. The first column shows the original images. The block sizes of other images range from 3 to 13 in increments of 2, from the left to the right column, respectively. (a) The azimuth angle is 0° and the elevation angle is 20°. (b) The azimuth angle is −50° and the elevation angle is −40°.

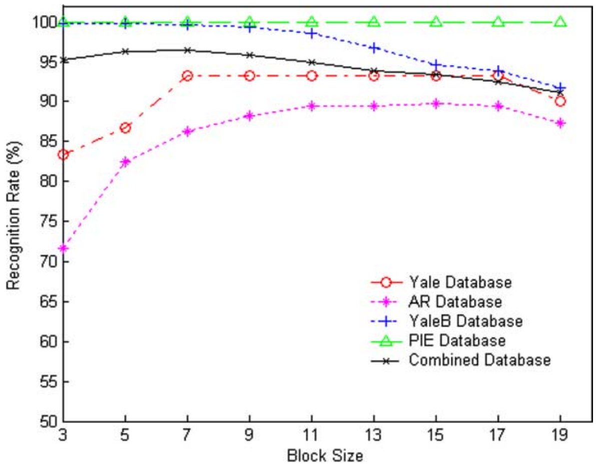

Fig. 4. Face recognition with different block sizes.

eigenfaces, and the mean image of the training set. In order to compare the recognition performances using the different databases, we used the combined database as the training set to generate a common set of eigenfaces, which are then used for image transformation and feature extraction. The number of eigenfaces used is 213. The Mahalanobis distance metric, which is a more suitable distance measure than the Euclidean distance metric for a standard PCA classifier (Yambor et al., 2002; Draper et al., 2003), is employed, and the nearest neighbor rule is then used to classify the face images. The experimental results are shown in Table 2. In the second row of Table 2, "None" means without using any preprocessing method to normalize/compensate the varying illuminations, and directly applying PCA for face recognition.

Table 2 shows that, with the different databases, our algorithm can achieve a better performance level than if no compensation/normalization scheme is used or if only the histogram equalization is used. The performance will slightly improve when the histogram equalization is used with the local normalization method; this shows that the global contrast enhancement can improve illumination compensation to a certain extent.

As the YaleB database is commonly used to evaluate the performance of illumination invariant face recognition, so we first compare our performance with other face recognition methods based on this database. Georghiades et al. (2000) proposed the individual illumination cone model and achieved 100% recognition rates, but the method requires seven images of each person to obtain the shape and albedo of a face. Lee et al. (2001) used a nine-point light source method to achieve a 99.1% recognition rate. However, the approach requires nine simulated images with different illumination variations for each person. Zhao et al. (2003) synthesized 45 images per person, which are adopted for training, and a 93.3% recognition rate was achieved. Liu et al. (2005) reported a 98.4% recognition rate. However, the iterative algorithm, which is used to restore the input image, is more computational than our method. All the above methods only consider the situation where the light source directions are within 75°, and so only 45 illumination models were used for testing. However, in our experiment, a total of 65 lighting conditions were tested. In our previous method (Xie and Lam, 2005a), the recognition rates are 99.5% and 96.4% when the respective eigenfaces and common eigenfaces are adopted for PCA method, respectively. The results are similar to those proposed in this paper. However, our previous method requires twenty feature points per image to determine the 2D shape of the input and to construct a shape-free texture, which is very difficult when the image is under varying or poor illumination. In (Liu et al., 2005), the recognition rate with the Yale database is reported to be 81.7%. Xie and Lam (2005a) have also tested their method based on the Yale database and AR database, and the results are 90.0% and 81.8%, respectively, when the respective eigenfaces are used, and 86.7% and 73.6%, respectively, when the common eigenfaces are adopted.

Compared to other methods, our proposed algorithm is much simpler. We neither require multiple images with different illumination variations as training, nor require the detection of important facial feature points to perform shape normalization. Our method is robust to illumination conditions and is computationally simple, which is important as a preprocessing method. Therefore, our method can also be used for other face recognition methods.

4.2.2. Face recognition using ICA

PCA can remove the pair-wise linear dependencies between pixels in an image, but high-order dependencies still exist in the joint distribution of the PCA coefficients. ICA (Bartlett et al., 2002; Liu and Wechsler, 2003; Déniz et al., 2003; Draper et al., 2003) can be considered a generalization of PCA, which can find some independent bases, namely independent components (ICs), by methods sensitive to high-order statistics. Then after image transformation and feature extraction, the ICA coefficients of an input computed based on the FastICA (Hurri et al., 2004) for ICA architecture II (Bartlett et al., 2002) are statistically independent. Compared to the eigenfaces, ICs retain more local information (Déniz et al., 2003). In this paper, we employed the FastICA to compute the ICs of a set of training images. FastICA provides rapid convergence and estimates the ICs by maximizing a measure of independence among the estimated original components (Déniz et al., 2003; Draper et al., 2003). The results in (Bartlett et al., 2002; Draper et al., 2003) show that ICA will have a better performance when the cosine similarity

Table 2
Face recognition results based on different databases using PCA

(%)	Yale	AR	YaleB	PIE	Combined
None	43.3	78.0	60.3	88.6	60.8
HE	50.0	81.0	63.3	96.8	68.4
LN	93.3	86.0	99.5	100.0	96.4
HE + LN	93.3	86.2	99.7	100.0	96.5

Table 3
Face recognition results based on different databases using ICA

(%)	Yale	AR	YaleB	PIE	Combined
None	40.0	77.4	65.6	95.1	64.8
HE	53.3	78.5	72.0	97.5	75.4
LN	83.3	82.4	98.1	100.0	90.6
HE + LN	86.7	82.6	99.8	100.0	94.5

measure is used. Therefore, we also adopt this similarity measure, which is defined as follows:

$$d(\boldsymbol{u},\boldsymbol{v}) = \cos\left(\frac{\sum_{i=1}^{k} u_i v_i}{\sqrt{\sum_{i=1}^{k} u_i^2}\cdot\sqrt{\sum_{i=1}^{k} v_i^2}}\right), \tag{12}$$

where u_i and v_i represent the ith element of two k-dimensional feature vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively. We also use the combined database as shown in Table 1 to produce the ICs, the number of ICs used being 214. The experimental results are shown in Table 3.

Comparing Tables 2 and 3, we can see that when the first two methods ('None' and HE) are used, ICA outperforms PCA in most of the cases. However, when our proposed LN method is employed with or without using the HE method, PCA outperforms ICA in most of the cases. As described in (Adini et al., 1997), uneven illuminations mainly affect the global components of a face image. Therefore, when the input image is under varying lighting conditions without any preprocessing method or when the HE method only is used for illumination normalization, ICA, which maintains more local, detailed information, performs better than PCA, which mainly considers the global structure of an input. This result coincides with the analysis in (Draper et al., 2003). When our LN method, which can effectively enhance the local structure of an image and reduce the global effect of the varying illumination, is used, more local and detailed texture will appear in the processed image. In this case, PCA can more effectively represent the more important structure of an image and reduce the effect of the noise enhanced by local normalization. Therefore, after the LN process, PCA outperforms ICA. In fact, the difference between these two methods is not large, especially for the YaleB database and the PIE database, where both methods can achieve a recognition rate near 100% (the Yale database is an exception, but its size is very small). We have also conducted some experiments in which the Euclidean distance metric is employed. For the combined database, the recognition rate without using any illumination normalization method is 62.4%, and the results using HE, LN and HE plus LN are 68.0%, 89.1% and 93.7%, respectively. These results are lower than those shown in Table 3, but the relative performances of these methods remain the same.

4.2.3. Face recognition using Gabor wavelets

The Gabor wavelets, whose kernels are similar to the response of the two-dimensional receptive field profiles of the mammalian simple cortical cell (Chui, 1992), exhibit the desirable characteristics of capturing salient visual properties such as spatial localization, orientation selectivity, and spatial frequency (Liu and Wechsler, 2003). In the spatial domain, a Gabor wavelet is a complex exponential modulated by a Gaussian function, which is defined as follows (Chui, 1992; Lee, 1996):

$$\psi_{\omega,\theta}(x,y) = \frac{1}{2\pi\sigma^2} \mathrm{e}^{-\left(\frac{(x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2}{2\sigma^2}\right)} \cdot \left[\mathrm{e}^{\mathrm{i}(\omega x\cos\theta+\omega y\sin\theta)} - \mathrm{e}^{-\frac{\omega^2\sigma^2}{2}}\right], \tag{13}$$

where x, y denote the pixel position in the spatial domain, ω is the radial center frequency, θ is the orientation of the Gabor wavelet, and σ is the standard deviation of the Gaussian function along the x- and y-axes, where $\sigma_x = \sigma_y = \sigma$ is assumed. The value of σ can be derived as follows (Lee, 1996):

$$\sigma = \kappa/\omega, \tag{14}$$

where $\kappa = \sqrt{2\ln 2}((2^\phi + 1)/(2^\phi - 1))$, and ϕ is the bandwidth in octaves. By selecting different center frequencies and orientations, we can obtain a family of Gabor kernels from (13), which can be used for representing an image. Given a gray-level image $f(x,y)$, the convolution of $f(x,y)$ and $\psi_{\omega,\theta}(x,y)$ is given as follows:

$$Y_{\omega,\theta}(x,y) = f(x,y) * \psi_{\omega,\theta}(x,y), \tag{15}$$

where $*$ denotes the convolution operator. The convolution can be computed efficiently by performing the fast Fourier transform (FFT), then point-by-point multiplications, and finally the inverse fast Fourier transform (IFFT). Concatenating the convolution outputs, we can obtain a one-dimensional Gabor representation of the input image,

$$\boldsymbol{Y}_{\omega,\theta} = [Y_{\omega,\theta}(0,0), Y_{\omega,\theta}(0,1), \ldots, Y_{\omega,\theta}(0,N_\mathrm{H}-1), Y_{\omega,\theta}(1,0), \ldots, Y_{\omega,\theta}(N_\mathrm{w}-1,N_\mathrm{H}-1)]^\mathrm{T}, \tag{16}$$

where N_w and N_H are the width and height of the image. In this paper, we only consider the magnitude of the Gabor representations, which can provide a measure of the local properties of an image (Lades et al., 1993) and is less sensitive to the lighting conditions (Shams and Malsburg, 2002) (for convenience, we also denote it as $\boldsymbol{Y}_{\omega,\theta}$). $\boldsymbol{Y}_{\omega,\theta}$ is normalized to have zero mean and unit variance distribution; and then the Gabor representations with different ω and θ are concatenated to form a high-dimensional vector as (17) and used for face recognition,

$$\boldsymbol{Y} = [\boldsymbol{Y}^\mathrm{T}_{\omega_1,\theta_1}, \boldsymbol{Y}^\mathrm{T}_{\omega_1,\theta_2}, \ldots, \boldsymbol{Y}^\mathrm{T}_{\omega_l,\theta_n}]^\mathrm{T}, \tag{17}$$

where T represents the transpose operation, and l and n are the numbers of center frequencies and orientations used. In our experiment, we select one center frequency, which is equal to $\pi/2$, and eight orientations from 0 to $7\pi/8$ in increments of $\pi/8$. The Euclidean distance metric is adopted and the nearest neighbor rule is used for classification. The experimental results are shown in Table 4.

Table 4
Face recognition results based on different databases using Gabor wavelets

(%)	Yale	AR	YaleB	PIE	Combined
None	63.3	90.9	86.7	99.9	86.1
HE	73.3	94.5	98.4	100.0	90.8
LN	100.0	98.3	99.4	100.0	98.4
HE + LN	100.0	98.6	99.5	100.0	98.7

Tables 2–4 demonstrate that, of the three feature extraction methods, Gabor wavelets can achieve the best performance. Especially for the PIE database, a 99.9% recognition rate can be obtained based on the original images. This is because Gabor wavelets can effectively abstract local and discriminating features, which are less sensitive to illumination variations. It is clear that applying our LN method can further increase the performance when using Gabor wavelets for face recognition based on different databases. Liu et al. (2005) also uses Gabor wavelets to extract features based on the restored images, and the recognition rate is 95.3% for the combined Yale database and YaleB database.

4.3. Computational complexity

We have proposed an efficient method of reducing the effect of varying illumination on face recognition. Suppose that the size of a normalized face is $M \times M$, and the block size used in the LN method is $N \times N$. The computational complexity for preprocessing an image using LN is $O(M^2N^2)$. All our experiments were conducted on a computer system with Pentium IV 2.4 GHz CPU and 512MB RAM. The average runtime of our algorithm to normalize the illumination of a face image in the AR database (363 face images) is about 6.2 ms, where M and N are equal to 64 and 7, respectively. As our method has a low complexity, it can also be applied to some real-time applications such as illumination normalization in video sequences.

5. Conclusions

In this paper, a novel and simple illumination normalization method for human face recognition under varying lighting conditions is proposed. A human face is treated as a combination of a sequence of small and flat facets. For each facet, the effect of the illumination can be modeled by a multiplicative term and an additive term. Therefore, a local normalization technique is applied to the image point by point. Local normalization can effectively and efficiently eliminate the effect of uneven illumination, and keep the local statistical properties of the processed image the same as for the corresponding image under normal lighting conditions. Then, the generated images, which are insensitive to illumination variations, are used for face recognition, and the performances are evaluated using different face recognition methods. Experimental results show that, with the use of PCA, ICA and Gabor wavelets for face recognition, the error rates can be reduced by 91.1%, 84.4% and 90.6%, respectively, based on the combined database when our illumination normalization algorithm is used.

A major advantage of our proposed method is that, for training, only one image per person under normal illumination is required; this is very important for real applications. In addition, there is no need to perform any facial feature detection and shape normalization, which can be very complicated when the lighting is uneven or complex. Furthermore, our method is computationally simple, can serve as a preprocessing technique and also combine with other methods for face recognition. In this paper, we only consider the situation where the human faces are frontal and have a neutral expression. For a practical face recognition application, various poses and expressions may combine with varying illuminations. If these effects are also considered, the overall recognition rates will be further improved.

Acknowledgment

This work was supported by a research grant from The Hong Kong Polytechnic University, Hong Kong.

Appendix A

The expectation and variance of $f'_P(x,y)$ within a facet W are equal to 0 and 1, respectively.

$$E(f'_P(x,y)) = E\left(\frac{f'(x,y) - E(f'(x,y))}{\mathrm{Var}(f'(x,y))}\right) = \frac{E(f'(x,y)) - E(f'(x,y))}{\mathrm{Var}(f'(x,y))} = 0, \quad (x,y) \in W.$$

$$\begin{aligned}\mathrm{Var}(f'_P(x,y)) &= \sqrt{\frac{\sum_{i=1}^{N}(f'_P(x_i,y_i) - E(f'_P(x,y)))^2}{N}} \\ &= \sqrt{\frac{\sum_{i=1}^{N}(f'_P(x_i,y_i))^2}{N}} \\ &= \sqrt{\frac{\sum_{i=1}^{N}\left(\frac{f'(x_i,y_i) - E(f'(x,y))}{\mathrm{Var}(f'(x,y))}\right)^2}{N}} \\ &= \frac{\sqrt{\frac{\sum_{i=1}^{N}(f'(x_i,y_i) - E(f'(x,y)))^2}{N}}}{\mathrm{Var}(f'(x,y))} \\ &= \frac{\mathrm{Var}(f'(x,y))}{\mathrm{Var}(f'(x,y))} = 1, \quad (x,y) \in W.\end{aligned}$$

References

Adini, Y., Moses, Y., Ullman, S., 1997. Face recognition: The problem of compensating for changes in illumination direction. IEEE Trans. Pattern Anal. Machine Intell. 19 (7), 721–732.

Ahlberg, J., 2001. CANDIDE-3—un updated parameterised face. Report No. LiTH-ISY-R-2326, January.

Bartlett, M.S., Movellan, J.R., Sejnowski, T.J., 2002. Face recognition by independent component analysis. IEEE Trans. Neural Networks 13 (6), 1450–1464.

Batur, A.U., Hayes, III, M.H., 2001. Linear subspaces for illumination robust face recognition. In: Proc. IEEE Conf. CVPR, Hawaii.

Belhumeur, P.N., Kriegman, D.J., 1996. What is the set of images of an object under all possible lighting conditions? In: Proc. IEEE Conf. CVPR, San Francisco.

Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. 19 (7), 711–720.

Bichsel, M., 1995. Illumination invariant object recognition. In: Proc. Internat. Conf. on Image Processing, Washington.

Chen, H.F., Belhumeur, P.N., Jacobs, D.W., 2000. In search of illumination invariants. In: Proc. IEEE Conf. CVPR, Hilton Head.

Chui, C.K., 1992. An Introduction to Wavelets. Academic Press, Boston.

Déniz, O., Castrillón, M., Hernández, M., 2003. Face recognition using independent component analysis and support vector machines. Pattern Recognition Lett. 24 (13), 2153–2157.

Draper, B.A., Baek, K., Bartlett, M.S., Beveridge, J.R., 2003. Recognizing faces with PCA and ICA. Computer Vision and Image Understanding 91 (1–2), 115–137.

Feng, G.C., Yuen, P.C., 2000. Recognition of head-and-shoulder face image using virtual frontal-view image. IEEE Trans. Systems Man Cybernet., Part A 30 (6), 871–882.

Georghiades, A.S., Kriegman, D.J., Belhumeur, P.N., 1998. Illumination cones for recognition under variable lighting faces. In: Proc. IEEE Conf. CVPR, Santa Barbara.

Georghiades, A.S., Belhumeur, P.N., Kriegman, D.J., 2000. From few to many: Generative models for recognition under variable pose and illumination. In: Proc. IEEE Conf. Face and Guesture Recognition, Grenoble.

Hallinan, P.W., 1994. A low-dimensional representation of human faces for arbitrary lighting conditions. In: Proc. IEEE Conf. CVPR, Seattle.

Hurri, J., Gävert, H., Särelä, J., Hyvärinen, A., 2004. Laboratory of Information and Computer Science in the Helsinki University of Technology. Available from: <http://www.cis.hut.fi/projects/ica/fastica/code/FastICA_2.1.tar.gz>.

Hwang, B.-W., Lee, S.-W., 2003. Reconstruction of partially damaged face images based on a morphable face model. IEEE Trans. Pattern Anal. Machine Intell. 25 (3), 365–372.

Ishiyama, R., Sakamoto, S., 2002. Geodesic illumination basis: Compensating for illumination variations in any pose for face recognition. In: Proc. 16th Internat. Conf. Pattern Recognition, Quebec.

Kirby, M., Sirovich, L., 1990. Application of the KL procedure for the characterization of human faces. IEEE Trans. Pattern Anal. Machine Intell. 12 (1), 103–108.

Lades, M., Vorbruggen, J.C., Buhmann, J., Lange, J., Malsburg, C.V.D., Wurtz, R.P., Konen, W., 1993. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput. 42 (3), 300–311.

Lam, K.M., Yan, H., 1996. Locating and extracting the eye in human face images. Pattern Recognition 29 (5), 771–779.

Lee, T.S., 1996. Image representation using 2D Gabor wavelets. IEEE Trans. Pattern Anal. Machine Intell. 18 (10), 959–971.

Lee, K.C., Ho, J., Kriegman, D., 2001. Nine points of light: Acquiring subspaces for face recognition under variable lighting. In: Proc. IEEE Conf. CVPR, pp. I-519–I-526.

Liu, C., Wechsler, H., 2003. Independent component analysis of Gabor features for face recognition. IEEE Trans. Neural Networks 14 (4), 919–928.

Liu, D., Lam, K.M., Shen, L.S., 2004. Optimal sampling of Gabor features for face recognition. Pattern Recognition Lett. 25 (2), 267–276.

Liu, D., Lam, K.M., Shen, L.S., 2005. Illumination invariant face recognition. Pattern Recognition 38 (10), 1705–1716.

Martinez, A.M., Benavente, R., 1998. The AR face database. CVC Technical Report #24, June.

Pentland, A., 2000. Looking at people: Sensing for ubiquitous and wearable computing. IEEE Trans. Pattern Anal. Machine Intell. 22 (1), 107–119.

Shams, L., Malsburg, C., 2002. The role of complex cells in object recognition. Vision Res. 42 (22), 2547–2554.

Shashua, A., Riklin-Raviv, T., 2001. The quotient image: Class based re-rendering and recognition with varying illuminations. IEEE Trans. Pattern Anal. Machine Intell. 23 (2), 129–139.

Sim, T., Baker, S., Bsat, M., 2002. The CMU pose, illumination, and expression (PIE) database. In: Proc. IEEE Conf. on Automatic Face and Gesture Recognition.

Sirovich, L., Kirby, M., 1987. Low-dimensional procedure for the characterization of human faces. J. Opt. Soc. Amer. 4, 519–524.

Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J. Cognitive Neurosci. 3, 71–86.

Wong, K.W., Lam, K.M., Siu, W.C., 2001. An efficient algorithm for human face detection and facial feature extraction under different conditions. Pattern Recognition 34 (10), 1993–2004.

Xie, X., Lam, K.M., 2005a. Face recognition under varying illumination based on a 2D face shape model. Pattern Recognition 38 (2), 221–230.

Xie, X., Lam, K.M., 2005b. An efficient method for face recognition under varying illumination. In: Proc. Internat. Symposium on Circuits and Systems, Kobe, Japan.

Yale University, 1997. Available from: <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.

Yale University, 2001. Available from: <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>.

Yambor, W., Draper, B., Beveridge, R., 2002. Analyzing PCA-based face recognition algorithms: Eigenvector selection and distance measures. In: Christensen, H., Phillips, J. (Eds.), Empirical Evaluation Methods in Computer Vision. World Scientific Press, Singapore.

Zhao, W., Chellappa, R., 2000. Illumination-insensitive face recognition using symmetric shape-from-shading. In: Proc. IEEE Conf. CVPR, Hilton Head.

Zhao, L., Yang, Y., 1999. Theoretical analysis of illumination in PCA-based vision systems. Pattern Recognition 32 (4), 547–564.

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