MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

15—4

Illumination Invariant Face Recognition Using Photometric Stereo

Seok Cheol Kee*, Kyoung Mu Lee**, and Sang Uk Lee*
*School of Electrical Engineering
Seoul National University, Seoul, Korea, {kee@diehard, sanguk@sting}.snu.ac.kr
**Department of Electronics and Electrical Engineering,
Hong-Ik University, Seoul, Korea, kmlee@wow.hongik.ac.kr

Abstract

In this paper, we propose an elegant approach for illumination invariant face recognition based on photometric stereo technique. The basic idea is to reconstruct the surface normal and the albedo of a face using photometric stereo images, and use them as the illumination independent model of the face. We also investigated the optimal light source directions for reconstructing accurate surface shape, and the robust estimation technique for the illumination direction of an input face image. The performance of the proposed algorithm is tested using 125 real face images of 25 person which are taken under 5 quite different illumination conditions, and achieved the success rate of more than 80%.

1 Introduction

Face recognition is a challenging subject because of dramatic variations between images of the same face. For example, facial expression, pose variation, hair style give rise to huge geometrical variations. In addition, illumination, aging, make-up, scale variation, background changes may cause misclassification. Thus, the face recognition system for the real-world applications is to recognize a face as independent as possible under these image variations. A recent survey [1] presents a detailed overview of relevant researches have been done for more than 20 years in the fields of psychophysics, neuroscience and engineering. Many psychophysical experiments [2] show that the human visual system can recognize faces against considerable variations between images of the same face. Among the image variations, we mainly focus on the problem of compensating for changes in illumination direction. Even though a human being has prominent ability to identify faces, he has difficulty in recognizing faces under severe illumination variation. The same person seen under varying illumination conditions usually appears quite different. Needless to say, the difference of a face induced by illumination condition is larger than that of individuals faces under the same illumination condition. So far few researches have been done for this illumination variation problem in face recognition including [6].

In this research, we propose a new method for illumination invariant face recognition based on photometric stereo. By constructing surface shape and albedo as the model of faces using three photometric images, and estimating the illumination direction of an input image, one can synthesize face images of the same illumination direction for the input face, so that robust illumination invariant matching can be performed.

2 Overview of the System

The whole procedure for the proposed face recognition algorithm is illustrated in Fig. 1. As the conventional recognition system, our system has two procedures: *registration* and *recognition*. For this purpose, two sets of face images; *training group* or *gallery*, and *testing group* or *prove* are constructed. Both procedures have the same preprocessing steps. First, face region is extracted using an ellipse-shaped mask by given two eye locations. And then, geometric and photometric normalization between photometric images as well as training and test images are necessarily.

In registration procedure, the shape model computed from training group forms the reference database. After the shape models of gallery are obtained, a generic shape model of face is constructed by averaging them. The generic shape model is used to estimate light direction in the recognition procedure.

In recognition procedure, the estimation method of light direction and matching problem are important steps for robust face recognition system. Once the shape models are obtained, the face image illuminated from arbitrary direction can be easily generated. Thus, by synthesizing the face image with the same illumination direction of an input face, we can match the two face images under the same illumination condition.

3 Face Shape Reconstruction using Photometric Stereo

We assume that a face has the Lambertian reflectance property such that the brightness is inde-

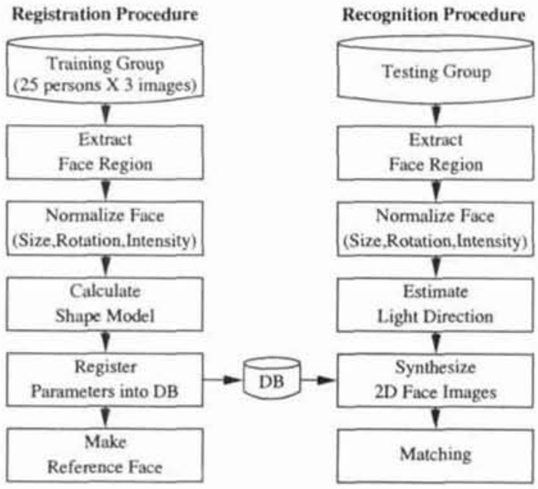


Figure 1: The overview of face recognition system

pendent on the viewing direction. Then, for a fixed light source direction, the image intensity at a point depends only on the surface normal $\mathbf{n}$ at that point [3] and this can be characterized by the image irradiance equation given by

$$I(x,y) = R(x,y) = \begin{cases} \rho(x,y)\mathbf{l}^T\mathbf{n}, & \mathbf{l}^T\mathbf{n} \geq 0, \\ 0, & \mathbf{l}^T\mathbf{n} < 0, \end{cases} \quad (1)$$

where R is the reflectance map, ρ is the albedo, and $\mathbf{l}$ is the unit vector of the illumination direction toward the point light source given by

$$\mathbf{l} \;=\; (\cos\tau\sin\sigma, \sin\tau\sin\sigma, \cos\sigma)^T. \quad (2)$$

The surface normal and the albedo of a Lambertian surface can be uniquely reconstructed using three images taken under different lighting conditions while maintaining the viewing direction fixed by the method known as the photometric stereo [3], [4].

Suppose that we have three photometric stereo images I_1, I_2, and I_3 obtained with the corresponding light source directions $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$, respectively. Then, we have three independent image irradiance equations for the images, which can be written in the matrix form as

$$\mathbf{I} = \rho\mathbf{L}\mathbf{n}, \quad (3)$$

where $\mathbf{I} = [I_1, I_2, I_3]^T$, and $\mathbf{L} = [\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3]^T$. If $\mathbf{l}_1$, $\mathbf{l}_2$, and $\mathbf{l}_3$ are linearly independent, $\mathbf{L}^{-1}$ exists. Then using the fact that $\mathbf{n}$ is unit normal, we can determine the albedo as

$$\rho = |\mathbf{L}^{-1}\mathbf{I}|, \quad (4)$$

and, in turn, the unit surface normal by

$$\mathbf{n} = \frac{1}{\rho}\mathbf{L}^{-1}\mathbf{I}. \quad (5)$$

4 Registration Procedure

4.1 Imaging geometry and optimal light source directions

We used the general imaging geometry model for acquiring photometric face images. The camera is positioned along the $\mathbf{Z}$-axis and the three light sources are located on a circle with the same slant angle σ, while different tilt angle τ's.

Selecting a set of appropriate photometric stereo images is very important for reconstructing accurate surface shape of a given object. Therefore, in order to choose the best photometric image set for a face, we first investigate the optimal illumination conditions for the face. The plaster *'Beethoven'* face is used for this test. For each slant angle of ($\sigma = 30°, 45°, 60°$), by varying the tilt angle of the first light source by the amount of 30° while maintaining the other light sources to be 120° apart in azimuth angle from each other, 12 sets of candidate basis images are obtained. The surface normals and albedo for each test set are determined using equations (4) and (5). Then the MAD (Mean Absolute Difference) error between real images illuminated from different directions and the corresponding synthesized ones are calculated to measure the reconstruction accuracy of each test set. We use three verifying images illuminated from $(\sigma,\tau) = (15°, 60°), (15°, 90°)$, and $(15°, 120°)$. The result are shown in Fig. 2 (a), and an example of verifying image, reconstructed image, and the error image is shown in Fig. 2 (b)-(d), respectively. In this test, in order to reduce the effect of cast shadow, we excluded the points that has the albedo value deviated far from the mean albedo, which are depicted as black pixels in Fig. 2(b)-(c).

Extensive experimental test with real plaster and synthesized face images reveals that the reconstruction error tends to be smaller when the slant angle (σ) relatively small and the tilt angle (τ) is arranged symmetrically. In this experiment, the light direction $(\sigma,\tau) = (30°, 90°), (30°, 210°)$, and $(30°, 330°)$ gives the smallest reconstruction error in any case, so that we select it as the optimal illumination direction.

4.2 Face region

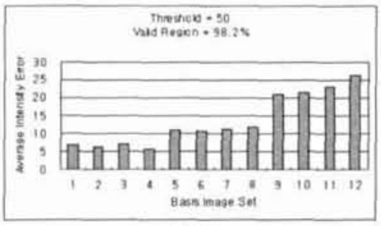
In order to consider only the facial region to be examined in a scene, we use an elliptic mask to extract facial area exclusively. Once locating the positions of eyes in a scene, the facial mask shape is designed and applied using the following elliptic function.

$$\frac{(x-c_x)^2}{a^2} + \frac{(y-c_y)^2}{b^2} = 1 \quad (6)$$

where

$$\mathbf{c} = (c_x, c_y), \qquad a = c_1 \cdot d_{eye}, \qquad b = c_2 \cdot d_{eye},$$

where $\mathbf{c}$ is center point between two eye locations, d_{eye} is distance between two eye locations, and c_1, c_2



(a)

(b) (c) (d)

Figure 2: (a) The average intensity errors of each basis set. (b) Target image obtained by a light source located at $(\sigma, \tau) = (15°, 120°)$. (c) Synthesized image generated with Set 4. (d) Error image magnified by 10 times.

are constants given empirically. Note that according to human races, the elliptic mask shape may be slightly different. In case of oriental Asian, by setting $c_1 = 1.00$, $c_2 = 1.43$, we can extract face region reliably in a scene.

4.3 Normalization

Photometric stereo assumes the correspondence between corresponding pixels in photometric stereo images. However, in real situation, although the camera direction is fixed, a face can be move slightly while taking pictures. This may cause mismatching between pixels, and result in inaccurate shape reconstruction. Therefore, it is necessarily to normalize the size and adjust the in-plane rotation of a face in each image. This can be accomplished by resizing and rotating the images using the eye positions as the anchor points. In addition, in order to compensate for the difference in illumination conditions between each image, we also normalize the intensity so that the maximum intensity value becomes unit for each image.

5 Recognition Procedure

5.1 Estimation of light direction

It is very important to estimate the illumination direction of an input face image correctly for accurate face matching. In this research, by constructing generic surface model of a face using the models in the gallery and utilizing the surface normal information of it, we estimate the illumination direction via least squared estimation method. Note that for i-th point in an input face image the following relation hold.

$$E_i = \frac{1}{\rho} I_i = \mathbf{n}_i^T \mathbf{l}, \tag{7}$$

where E_i is the normalized intensity of the i-th point in the input face image, $\mathbf{n}_i = [n_x^i, n_y^i, n_z^i]^T$ is the surface normal of the corresponding point in the generic model, and $\mathbf{l} = [l_x, l_y, l_z]^T$ is the illumination direction vector. If we are considering m pixels in a face region, we have m linearly independent equations, which can be written in following matrix form as

$$\begin{aligned} \mathbf{E} &= [\ E_1, E_2, \cdots, E_{m-1}, E_m\ ]^T \qquad (8) \\ &= \begin{bmatrix} n_x^1 & n_y^1 & n_z^1 \\ \vdots & \vdots & \vdots \\ n_x^m & n_y^m & n_z^m \end{bmatrix} \begin{bmatrix} l_x \\ l_y \\ l_z \end{bmatrix} = \mathbf{Nl}, \end{aligned}$$

where $\mathbf{E}$ is the normalized intensity matrix of the input face image, $\mathbf{N}$ is surface normal matrix of the generic face model. Then the illumination direction can be determined by least square estimation as

$$\mathbf{l} = (\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{E}. \tag{9}$$

Note that the relatively flat face regions such as forehead and cheeks usually have not enough intensity variations by the change of the illumination direction. Thus, in order to enhance the accuracy of the estimation, we consider only the region around the nose where the illumination sensitivity is relatively good, rather than whole face for calculating (9).

5.2 Matching

Once the illumination direction of a input face image is obtained, the face images of all the gallery models for the same illumination condition can be synthesized using (1). Then by measuring the similarity between the input face and the synthesized model face images, illumination invariant face matching can be accomplished. The dissimilarity measure used in this experiment is MAD of the intensities between the input and synthesized model face given by

$$D = \frac{1}{N} \sum_{(x,y) \in R} |I_i(x,y) - I_m(x,y)|, \tag{10}$$

where R is face region, and N is the number of pixel within face region, and I_i, I_m are the intensities of the input image and a synthesized image, respectively. By ordering the model faces according to the dissimilarity in increasing order, we can establish a registered face list for the given input face. Note that usually the recognition rate is defined by the percentage that the input face ranks top in the registered face list. However, since this does not represent the potential ability of a recognition scheme well, we employ the *cumulative recognition rate* [5]. The cumulative recognition rate represents the probability of the input face being matched to the first top n faces on the registered list.

6 Experimental Results

All face images were acquired in a darkroom isolated from unwanted extra light sources. The 640 x 480, 8-bit grey level images were taken using 3CCD digital camcorder interfaced to a PC.

In this experiment, images of 25 distinct persons, 20 males and 5 females, were taken in frontal viewing direction with different image size, and without any restriction about distinctive features such as glasses, make-up, or hair-style. Two groups of DB are constructed: One is the training group for basis images of each person, which are taken under the optimal illumination directions discussed in Section 4.1. The other is the testing group for evaluating the performance of the proposed algorithm. 5 different images for each person taken under very different illumination directions from that of the basis images are used for the test. Fig. 3 shows some face images of the testing group.

To investigate the performance variation due to the accuracy of the illumination direction estimation, the results for the ground truth and for 2 different light direction estimation method which uses the whole face region and the nose region, respectively, are evaluated. Table 1 represents the average estimation error (in radian angle) of the illumination direction estimation methods, and Fig. 4 shows the cumulative recognition rates for them. The result shows that when the nose region are used for the estimation, the proposed algorithm can achieve a top 1 recognition rate of 80% even under a severely changing illumination condition as shown in Fig. 3. It also shows that the cumulative recognition rate can increase to 93% if $n = 10$. While, note that when the whole face region is engaged in estimating the light direction, the recognition rate decreases about 10% and the estimation error also increases 0.09 radian. Therefore, we conclude that selection of effective region for estimating the illumination direction is very critical for accurate face recognition. Note also that if we estimate illumination direction exactly, the recognition rate will increase up to 87%, and we can achieve potential performance of the excellent cumulative recognition rate of 96%.

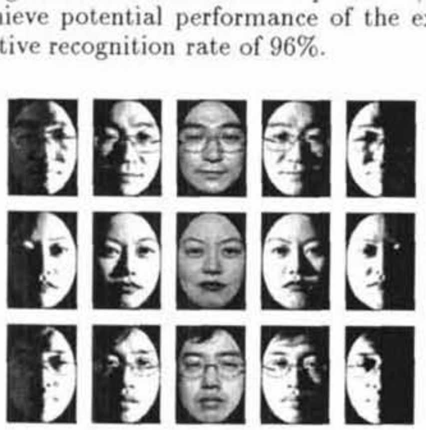


Figure 3: Some face images of testing group. All images were shown after normalization steps.

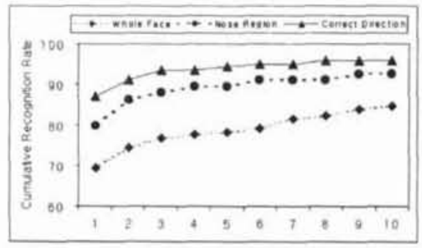


Figure 4: The cumulative recognition rates at different light direction estimation methods.

Table 1: Light direction estimation errors

Estimation Method	Average Error
Whole face	0.21
Nose region	0.12
Correct direction	0.00

7 Conclusions

We propose a new approach for illumination invariant face recognition, in which the shape information of a face is determined by the photometric stereo method, and used as the illumination independent model of the face. The optimal illumination directions for three basis images which insure the accurate shape reconstruction is also investigated. Experimental results on real face images show that the proposed approach has a great potential for the robust face recognition even when the lighting condition changes significantly.

References

[1] R. Chellappa, C. L. Wilson and S. Sirohey, "*Human and Machine Recognition of Faces: A Survey,*" Proc. IEEE, vol. 83, no. 5, pp. 705-740, 1995.

[2] V. Bruce and M. Burton, *Processing Images of Faces.* Norwood, New Jersey: Ablex Publishing Corporation, 1992.

[3] R. J. Woodham, "*Photometric Method for Determining Surface Orientation from Multiple Images,*" Optical Engineering, vol. 19, No. 1, pp. 139-144, 1980.

[4] K. Ikeuchi, "*Determining Surface Orientations of Specular Surfaces by Using the Photometric Stereo Method,*" IEEE Trans. PAMI, vol. 13, No. 6, pp. 661-669, July 1981.

[5] K. Lam and H. Yan, "*An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View,*" IEEE Trans. PAMI, vol. 20, No. 7, pp. 673-686, July 1998.

[6] A. S. Georghiades, D. J. Kriegman, and P. N. Belhumeur, "*Illumination Cones for Recogniton Under Variable Lighting : Faces,*" In Proc. IEEE Conf. on CVPR, pp. 52-58, 1998.