MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

3—26

A Face Recognition Algorithm Robust Against Illumination Variations Using 3-Dimensional Face Shape

Hitoshi Imaoka
System Devices and Fundamental Research
NEC Corporation

Shizuo Sakamoto
Computer & Communication Media Research
NEC Corporation

Abstract

The authors propose a face recognition algorithm robust against the illumination variations, using *an average face shape*. The objective of this paper is two-fold: (1) to propose a new algorithm to estimate one of the illumination-invariant features, i.e., the surface reflectance from single appearance image, and (2) to apply this algorithm to images with various illumination conditions. In (1), we take advantage of *the average face shape* for local area of face so as to suppress the deviation of shape from the individual one. The correct recognition rate (97.0%) for images of 41 persons with 5 illumination conditions was much higher than that by the edge-based matching method (82.9%).

1 Introduction

Machine recognition of face is very useful technique because this has many applications such as the personal identification of driver's license/passport image, the robot vision[1], the crowd surveillance and the bank/store security[2]. One of difficulties in the face recognition is the variability of an object's appearance. Robustness for this variability is crucial for the practical use.

In the real-world scene, there are various kinds of variabilities such as view angle, illumination condition, face expression and aged deterioration. In the previous study[3], the authors proposed a recognition method robust against the view angle, and succeeded in obtaining a practical recognition rate. The core of this method is to generate another view of face from one appearance image by using an *average face shape*. The present paper focuses on the robustness for illumination variations.

Here, we describe the condition of illumination. For example, in the case of the narrow spotlights whose intensities is varying for each local area of face, it is possible to generate a great variety of images. From this reason, it is clear that considering for all possible illumination is no meaning. We limit to combination of light sources whose intensities are stable within the scale of face size.

The existing methods for handing the illumination variabilities are roughly classified into three approaches[4,5]: 1) the universal approach[6,7], 2) the model-based approach[8-10] and 3) the class-based approach[11,12]. The universal approach uses a illumination insensitive feature which is common to all images independent of the specific set of objects to be recognized. Typical example of this approach is edge-based matching method. Discontinuities in the albedo on the surface of the object generate edges in images. Hence, the edges tend not to be influenced by the change of illumination. However, as discussed in ref.[6], the ability to recognize face is not practical when the difference of illumination condition is large. The model-based approach uses information specific to the object, such as 3D face shape and the surface reflectance. This approach compensates for image variations extremely, but it is difficult to collect the information for each object. At class-based approach, the variations by changing the illumination are handled by the general information of object of class. In this approach, a feature insensitive illumination is derived from incorporate information about object. An example of this approach is shape from shading[11]. In these approaches, the authors in ref.[5] conclude that the class-based approach is advantageous in novel viewpoint and illumination conditions because it is more specific than universal approach and more general than model-based approach.

The present paper proposes a new algorithm to estimate one of the illumination-invariant features, i.e., the surface reflectance from single appearance image, using the *average face shape*. This method belongs to class-based approach and has advantages: a) the system using this method has wide applications, because the registration data is only single image for each person, and b) by combining the method proposed in [3], it is possible to attain the robustness for view and illumination variations. Furthermore, as the relational study, Volker et.al. proposed the robust method for illumination and viewpoint using morphable face model[13] and Belhumeur et.al.

*Address: 34 Miyukigaoka, Tsukuba, Ibaraki 305-8501 Japan. E-mail: h-maoka@cb.jp.nec.com
4-1-1 Miyazaki, Miyamae-ku, Kawasaki, Kanagawa 216-8555 Japan. E-mail: s-sakamoto@bu.jp.nec.com

show the basis set of images of an object created under all possible illumination conditions[10]. Combination of both methods make possible to recognize object under all possible condition, but much computational complexity needs in order to optimize free variables subject to the illumination condition and face shape.

Section 2 shows an algorithm to estimate the surface reflectance, and section 3 demonstrates the experimental result of the individual face identification. Section 4 describes the concluding remarks and the future problem.

2 Estimation of surface reflectance

We start from the Phong's illumination model:

$$\begin{aligned} I_i &= a_i I_{d,i} + b I_{s,i} \\ &= a_i \max(\mathbf{L}_i \cdot \mathbf{N}_i, 0) + b \max((\mathbf{L}'_i \cdot \mathbf{V})^n, 0), \end{aligned} \tag{2.1}$$

where $i(=1,2,\cdots,N)$ is the number of pixel, $I_i = (I_{R,i}, I_{G,i}, I_{B,i})^t$ is the image intensity of i-th pixel, $a_i = (a_{R,i}, a_{G,i}, a_{B,i})^t$ is the surface reflectance, $b = (b_R, b_G, b_B)^t$ is the light source color, $I_{d,i}$ is the diffused reflection component, $I_{s,i}$ is the specular reflection component, $\mathbf{L}_i$ is the light source direction with intensity, $\mathbf{N}_i$ is the surface normal, $\mathbf{L}'_i = 2(\mathbf{L}_i \cdot \mathbf{N}_i)\mathbf{N}_i - \mathbf{L}_i$ is the direction of reflected ray of $\mathbf{L}_i$, $\mathbf{V}$ is the direction of camera position and n is the constant to control the angular distribution of the specular reflection. Subscripts (R, G, B) denote primary color of image. In the above parameters, the image intensity I_i and the surface normal $\mathbf{N}_i$ are known, and other parameters are unknown.

An algorithm to estimate the surface reflectance is composed of the following three steps: (i) separation of specular reflection component from the face image, (ii) estimation of illumination direction, and (iii) calculation of surface reflectance. The details of these steps are given as follows.

In step (i), we assume that the surface reflectance a_i is independent of pixel i. Then, eq.(2.1) is written by

$$I_i = a I_{d,i} + b I_{s,i}. \tag{2.2}$$

Since a and b are unknown, we regard a as a flesh color and b as a white color ($b_R = b_G = b_B$). Here, the flesh color is decided by averaging the color of sample pixel in skin area. By the least square method,

$$(I_{d,i}, I_{s,i})^t = (A^t A)^{-1} A^t I_i, \tag{2.3}$$

where $A = (a, b)$ and the subscripts t and -1 denote the transposed matrix and the inverse matrix, respectively.

Step (ii) is carried out every local area of face. The reasons are: 1) to suppress the deviation of the *average face shape* from individual one and, 2) to consider the case such that the light source direction varies depending on the face local area. The diffused reflection for local area is described by

$$I_{d,i} = a_i \max(\mathbf{L}_{p_i} \cdot \mathbf{N}_i, 0), \tag{2.4}$$

where $\mathbf{L}_p$ is the illumination direction at pixel p, i is the pixel that contains area Ω within radius r of pixel p in the face image plane. The radius r is taken so that skin occupies almost of area Ω. Assuming that i) $a_i = 1$ for any i and ii) $\max(\mathbf{L}_{p_i} \cdot \mathbf{N}_i, 0) \approx \mathbf{L}_{p_i} \cdot \mathbf{N}_i$, we obtain

$$\hat{\mathbf{L}}_p = \mathbf{I}_{d,\Omega} \mathbf{N}^t_\Omega (\mathbf{N}_\Omega \mathbf{N}^t_\Omega)^{-1}, \tag{2.5}$$

where $\mathbf{I}_{d,\Omega} = (I_{d,1}, I_{d,2}, \cdots, I_{d,N_\Omega})$, $\mathbf{N}_{d,\Omega} = (\mathbf{N}_{d,1}, \mathbf{N}_{d,2}, \cdots, \mathbf{N}_{d,N_\Omega})$, N_Ω the number of pixel contained in Ω. Since solution in eq.(2.5) is valid under above two assumptions, we calculate eq.(2.5) iteratively by eliminating pixels i with $(\hat{\mathbf{L}}_p \cdot \mathbf{N}^t_{d,i}) < 0$ or $\left| I_{d,i} - \hat{\mathbf{L}}_p \cdot \mathbf{N}^t_{d,i} \right| > \varepsilon$ from Ω. Here ε is the constant.

Step (iii) calculates the surface reflectance by following equation:

$$a_i = (I_i - b I_{s,i}) / (\mathbf{L}_i \cdot \mathbf{N}_i), \tag{2.6}$$

Since the coefficient of surface reflectance is not uniquely decided, a_i in eq.(2.6) is normalized to $(1,1,1)^t$ for skin area.

Figure 1 shows the surface reflectance obtained from above three steps. Figures 1(a), (b), (c) and (d) denote an original face image, the diffused reflection component, the specular reflection component and the surface reflectance. Here, the black part in the tip of the nose in Figs.1(b) and (d) express pixels with halation in the original image. From Fig.1(c), it is found that a specular part of the right side of forehead is extracted accurately.

3 Experimental results

The proposed algorithm was applied to images with 5 illuminations conditions for 41 persons. Figure 2 shows examples of these images for one person. Figure 2(a) was "shadow-less image" that was generated by 4 light sources. Figures 2(b) to (e) were shading images shot under one light source whose directions were (−30°,35°), (0°,35°), (30°,35°), (60°,35°) for the horizontal and upper directions of face, respectively. The distances to the light sources were about 2 *meters*. We call the illumination conditions corresponding to Fig.2(a) to (e), *L0 to L4*, respectively.

As shown in Fig.3, experiments were composed of three processes: normalization of face area, the estimation of surface reflectance and matching of features. One registration image for each person was accumulated in the face template database. To investigate the difference by the matching features, four experiments were performed: 1) matching by intensity of image, 2) matching by the edge feature[6], one of the appearance-based approaches, 3) matching by the surface reflectance, the proposed method, 4) matching by the surface reflectance, where the individual face shape is used instead of the *average face shape*. For all experiments, a normalized cross-correlation method was used for comparison. Among the registration images, the person with the high score of similarity of matching feature was recognized as the identical one. The results of correct rate were shown in Table 1. Average rates were 1) 62.6%, 2) 82.9%, 3) 97.0% and 4) 100.0%.

4 Concluding Remarks

We considered the face recognition algorithm robust against the illumination variations. Assuming the Phong's illumination model and the average face shape, we obtained the surface reflectance from single appearance image. From the identification experiments, we found that

- the correct rate of the proposed method is superior to that of the edge-based method in spite of the same registration data
- the estimated error of surface reflectance is mainly occurred from the deviation of face shape from the experiment 3) and 4).

In the future, we will consider the robustness for other variations such as passage of age, using the neural network approach.

References

[1] Personal Robot: R100; http://www.incx.nec.co.jp/robot/indexj.html
[2] R. Chellappa, C.L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces," *Proceedings of IEEE*, Vol.83, No.5, pp.705-740, 1995.
[3] H.Imaoka and S. Sakamoto, "Pose-Independent Face Recognition Method," *Technical Report of IEICE PRMU*, 99-26, pp.51-58, 1999.
[4] Y. Adini, Y. Moses, and S. Ullman, "Face Recognition: The Problem of Compensating for Changes in Illumination Direction," *IEEE Transactions of Pattern Analysis and Machine Intelligence* Vol.19, No.7, pp.721-732 1997.
[5] Y. Moses and S. Ullman, "Generalization to Novel Views: Universal, Class-based, and Model-based Processing," *International Journal of Computer Vision*, Vol.29, pp.233-253, 1998
[6] R.Brunelli and T.Possio, "Face recognition: Feature versus templates," *IEEE Transactions of Pattern Analysis and Machine Intelligent*, Vol.15, No.10, pp.1042-1052, 1993.
[7] L. Wiskott, J-M. Fellous and N. Kruer and C. Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.775-779, 1997.
[8] R.Ishiyama, S.Sakamoto and J.Tajima, "A New Face-Recognition system with robustness against illumination changes," *to be appear on proceedings of MVA 2000*, 2000.
[9] A.S.Georghiades, D.J.Kriegman and P.N.Belhumeur, "Illumination cones for recognition under variable lighting: faces," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.52-58, 1998.
[10] P.N. Belhumeur and D.J.Kriegman, "What Is the Set of Images of an Object under All Possible Illumination Conditions?," *International Journal of Computer Vision*, Vol.28, No.3, pp.245-260, 1998.
[11] B.K. Horn, "Understanding Image Intensities," *Artificial Intelligence*, Vol.8, pp.201-231 1997.
[12] S.C. Kee, K.M. Lee and S.U. Lee, "Illumination Invariant Face Recognition Using Photometric Stereo," *Proceedings of IAPR Workshop on Machine Vision Applications*, pp.581-584, 1998.
[13] B. Volker and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," *Siggraph Conference Proceedings*, pp.187-194, 1999.

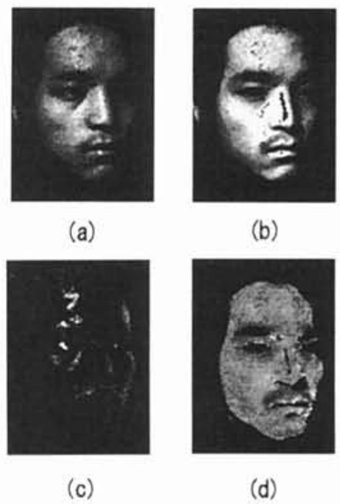

Fig.1 (a) the original face image, (b) the diffused reflection component, (c) the specular reflection component, (d) the surface reflectance.

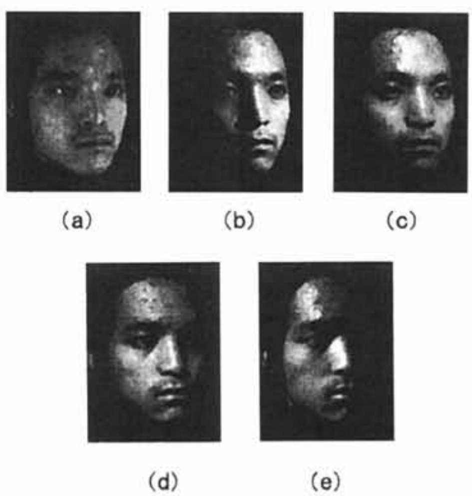

Fig.2 Examples of face images used in the experiment. (a) the registration image, (b)~(e) the input images

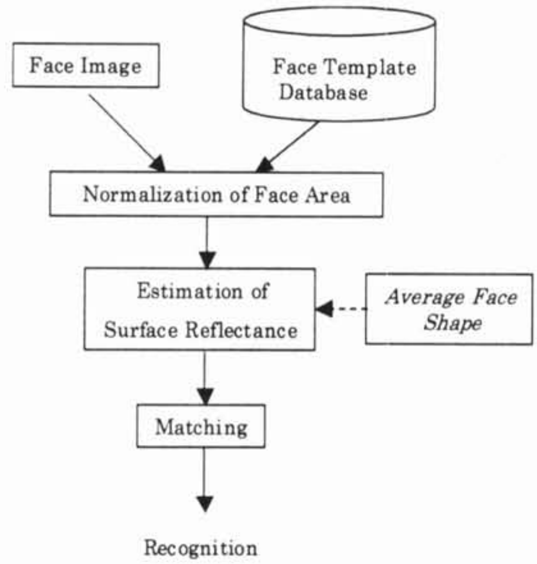

Fig.3 Scheme of face recognition

Table 1 Correct rate for each experiment (%).

Registration image vs Input image	No.1	No.2	No.3	No.4
L0-L1	73.2	97.6	100.0	100.0
L0-L2	39.0	80.5	100.0	100.0
L0-L3	70.7	85.4	100.0	100.0
L0-L4	68.3	65.9	95.1	100.0
L1-L2	95.1	100.0	100.0	100.0
L1-L3	26.8	68.3	85.4	100.0
L2-L3	68.3	100.0	100.0	100.0
L2-L4	22.0	48.8	92.7	100.0
L3-L4	100.0	100.0	100.0	100.0
average	62.6	82.9	97.0	100.0