Recognizing Face Images under Different Lighting Conditions

Haitao Wang, Yangsheng Wang
National Laboratory of Pattern Recognition,
Institute of Automation, Chinese Academy of Sciences
E-mail: htwang@nlpr.ia.ac.cn

Abstract

In this paper, we present a new algorithm for recognizing face images under different lighting conditions. Instead of the commonly used methods of modeling the 3D shape or finding some illumination invariant, the basic idea of our algorithm is to create new synthetic images by simulating lighting effects of the input image for each candidate image in the database, and then make correlation checking between the input image and the resulting synthesized images. With our method, only a single image is needed for each face in database, which significantly decreases the size of the database. Theory analysis and experiment results show that this algorithm can strikingly enhance the recognition rate and can be used for real time applications.

1. Introduction

In the past decade, the research of face recognition was focused mainly on pose changes under the same or similar illumination conditions. As opposed to this, few efforts have been made for varied lighting conditions. In real applications, however, differences due to the lighting changes may significantly exceed the differences due to different identities and slight changes in the lighting conditions, and, hence, will cause large changes in face's appearance. This is the reason why illumination is listed in the table of the main difficulties for face recognition.

There are three commonly used approaches to dealing with image variations that are due to illumination changes: the 3D method, the illumination invariant method, and the appearance based method. The idea of the 3D method is to use 2D images of a face to model the shape and albedo of the face. SFS [2](shape form shading) and photometric stereo [3] are two main tools for this approach. This method needs too many assumptions, and the corresponding modeling process is computationally expensive, which limits its use. The illumination invariant approach is to find some illumination invariant or some illumination insensitive representations of image changes under different illumination condition. However, the experiment results [1] show that many typical compensation algorithms for illumination changes are not invariant or insensitive to illumination at all. While the appearance based approach [6] is to cope with illumination changes by sampling all the illumination conditions for the same face. Unfortunately, this would lead to exponential growth of size of the image database.

Recently, two new methods, the illumination cone [5] and the quotient image [4], are presented. Both of them have shown their potencies in face recognition. It is claimed that the illumination cone algorithm can achieve the highest recognition rate under different illumination conditions, and can recognize face image with large shadows. In order to reach the goal, however, seven distinct images of the same face [5] are needed to estimate its 3D shape and albedo map up to a generalized bas-relief (GBR)[10] transformation. After the 3D reconstruction, a large number of images of extreme rays are rendered to form a convex cone for a fixed pose in the image subspace. In the formation of illumination cone, shadows (attached and cast) should be detected to avoid these illegal data, which do not satisfy the Lambertian assumption. Another disadvantage of the method is that, depending on the input data, it may be very time-consuming [5]. Furthermore, some related problems are left untouched. For instance: do the illumination cones of different faces with same pose or the illumination cones of different faces with different poses have intersections?

If only an extremely illuminated frontal face image and a database with only a unique image of each face are given, no existing algorithm can recover its original effects (non-shadow frontal illuminated image). Instead of commonly used methods of compensating for the illumination changes, our algorithm estimates the illumination direction of the test image and simulates the same illumination effects for all the images in the database. The recognition is made between the test image and synthetic images. The experiments show that our algorithm can strikingly improve the recognition rate.

Because our algorithm is inspired by the quotient image, we briefly discuss the quotient image and some of its limitations before presenting our method in section 3.

2. The Quotient Image Algorithm and its Limitations

The quotient image, as Shashua [4] pointed out, has the advantage of simple and clean theoretical foundation. It is non-iterative and has limited calculation complexity.

This algorithm assumes an ideal Class of Objects, which has the same shape but differs in the surface albedo function. The image space of such a class is represented by

$$\rho_i(x, y) n(x, y)^T s_j , \qquad (1)$$

where ρ_i is the albedo (surface texture) of object I of the class, $n(x,y)^T$ is the surface normal (shape) of the object (the same for all objects of the class), and s_j is the point light source direction, which can vary arbitrarily. The quotient image Q_y of object y against object a is defined by

$$Q_y(u, v) = \frac{\rho_y(u, v)}{\rho_a(u, v)} = \frac{\rho_y(u, v) n(u, v)^T s_y}{\rho_a(u, v) n(u, v)^T s_y}$$

$$= \frac{I_y}{\rho_a(u, v) n(u, v)^T \sum_j x_j s_j}$$

$$= \frac{I_y}{\sum_j I_j x_j}$$

$$j=1,2,3 \qquad (2)$$

where u and v range over the image, I_y is an image of object y with the illumination direction s_y, and I_j is an image of object a with non-collinear illumination direction s_j.

Using the $3N$ matrices $A_1, A_2, \ldots, A_N$ as the bootstrap set, whose columns are three face images of object $i \in (1, 2, \ldots, N)$ with non-collinear illumination direction, the problem is to estimate the illumination direction of s_y by $x = (x_1, x_2, x_3)$, which satisfies

$$s_y = \sum_j x_j s_j \qquad (3)$$

By minimizing an energy function, the illumination direction can be estimated.

From the definition (2), the quotient image Q_y is illumination free and therefore is lighting invariant, but according to our experiments, it is not. The synthesis results by quotient image method are shown in Figure 1. Fig. 1(a) is the bootstrap set and the first column of Fig.1(b) is the input images. The corresponding quotient images are show in the second column of Fig.1 (b) and synthetic images are from column 3 to 6 of Fig.1 (b).

From the first two rows of Figure 1(b), this algorithm shows its ability to synthesize new images with shadows, while this algorithm fails to synthesize reasonable images in the last two row of Fig. 1(b) because of the large shadows in the input images. In addition, all the shadow patterns of the synthetic images are in fact a linear combination of the bootstrap set and no shadow pattern like the last row input image in Fig. 1(b) can be synthesized by the bootstrap Fig. 1(a). Moreover, the same person in Fig. 1(b) row 3 and row 4 has different quotient images.

It is clear that the quotient image is not invariant under different illumination condition, especially for the input image with large shadows. This results can be deduced from equation (1), (2) and (3).

In most cases, the equation (1) should be written as

$$I(x, y) = \max(0, \rho_i(x, y) n(x, y)^T s_j) \quad (4).$$

If there are attached shadows, the product of $n(x, y)^T s_j$ is less than zero and, then, the value of $I(x,y)$ should be set to zero. By equation (2) and (3), they assume that the linear combination of bootstrap set images I_j can simulate the illumination direction of I_y. This assumption can be valid only when the object is perfectly convex and there will be no shadows. Because of the concavity of face, there are always some shadows in the face images and it is impossible to simulate all the image illumination effects by only three kinds of images.

(a) The bootstrap set

(b) Synthetic results of quotient image algorithm

Figure 1 the bootstrap set and synthesis results

3. Our New Method

Though the quotient image method can not provide an illumination invariant representation, it does show a novel approach for image synthesis. The quotient image, in fact an image ratio, can be used to map one image to another.

Given only one non-frontally illuminated face image, it impossible to recover it to its frontally illuminated image unless the face's 3D shape is known, such as in the case of illumination cone. But we can estimate its illumination pattern and transform all the images in the database into the similar pattern by quotient image technique. To overcome problems brought about by the shadow, a larger bootstrap set with systematically sampled face images is adopted.

Let $B = [b_1, b_2, \ldots, b_N]$ is bootstrap set, which is $M \times N$ matrix (M is the pixel number of a image), and y_s is a face image with mx1 format.

To estimate the illumination effects, an energy function $f(x)$ is still needed to minimize the errors.

$$f(x) = (y_s - Bx)^T (y_s - Bx) \quad (5)$$

We can get the minimum of *f(x)* by simple least squares technique.

$$0 = \frac{\partial f}{\partial x} = -B^T y_s + B^T Bx$$
$$x = (B^T B)^{-1} B^T y_s \quad (6)$$

We call x illumination factor because it no longer denotes illumination direction as in equation (3). In equation (6), x is the coefficient of the illumination effects. The minimization process in equation (5) will make the Bx simulate the illumination effects of y_s. Before calculating the illumination factor x in equation (6), an SVD transformation from B to B^* is done to make the $(B^T B)$ invertible.

$$B = SVD(B) = USV$$
$$B^* = U^* S^* V^*$$

where U is $M \times M$ matrix, V is $N \times N$ matrix, S is diagonal matrix, S^* is a diagonal $P \times P$ matrix with non-zero elements of S, and U^*, V^* are its corresponding P rows and P columns.

We have the quotient images for each face in the database by the ratio of Yi and Bx_i,

$$Q_i = \frac{Y_i}{Bx_i}, \qquad i = 1,2,...,K \quad (7)$$

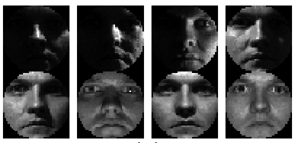
where K is the number of image in the database, and x_i can be estimated by equation (6).

If y_t is the input image for recognition and its estimated illumination factor is $x=[x_1,x_2,x_3]^t$, we have

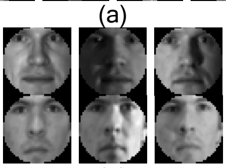
$$Y_{syn_i} = Q_i \otimes \sum_j B_j x_j \quad (8),$$

where $\otimes$ denotes the Cartesian product (pixel by pixel multiplication). Then the recognition is done between the input image Y_t and the Y_{syni}.

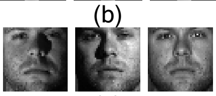
4. Experiments and Discussions



(a)



(b)



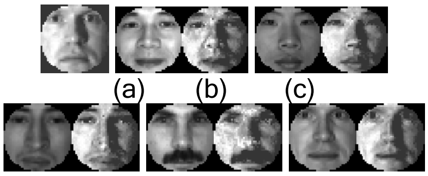
(c)

(d)

Figure 2 (a) part of our bootstrap set, (b) part of our test and training images, (c) the Yale B set 1 (d) the Yale B set 2

Our bootstrap set contains the 65 images of a person with frontal view under different illumination directions and the ambient illumination direction from Yale Face Database B [9]. And the recognition is testified in the Yale Face Database, Yale Face Database B and the MIT face database [6]. The part of bootstrap set and face database are shown in Figure 2.

Only frontal pose is assumed for all the three database and the face images are manually segmented and masked. The training set contents frontal illuminated images and the test set is composed of image with large shadows. Because the illumination direction variation in Yale B database is so great that we extract two Sets from the database: Set 1 with either azimuth or elevation less than 35° and Set 2 with either azimuth or elevation more than 35° but less than 70°.



(a) (b) (c)

(d) (e) (f)

Figure 3 (a) the input image, (b)(c)(d)(e)(f) the face image in training set and its corresponding synthetic image according to illumination effects of input (a)

Figure 3 shows the recognition process. Given the input image Fig.3 (a), our new algorithm will transform all the images in training set into the same illumination pattern as input Fig.3 (a) and then the recognition is carried out between the synthesis images and input image.

Table 1 the recognition results

Test SET	Our algorithm	Direct correlation	Quotient image
(1) MIT database	84.4%	75%	84.4%
(2) Yale database	76.7%	33.3%	76.7%
(3) Yale B Set 1	100%	96%	97.0%
(4) Yale B Set 2	83.8%	70.5%	31.9%

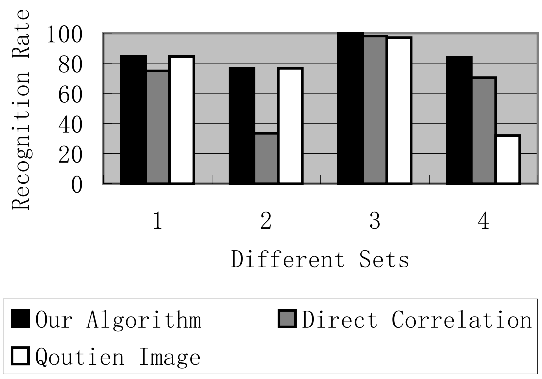


Figure 4 The results of recognition

To verify the effectiveness of our algorithm, we use the same classifier for our algorithm, direct correlation between input image and the database images and quotient image method. The classifier used is the nearest neighbor classifier, which employs correlation coefficient as the measurement for classification. The recognition results are shown in Table 1 and Figure 4.

It is obvious from the results of our algorithm and the direct correlation method that the recognition rate can be remarkably improved except in Yale B Set 1, in which the illumination variants is relatively small. The results of Yale B Set 1 also demonstrate that our algorithm do not affect the recognition rate under normal nearly frontal illumination conditions. Compared with quotient image method, our algorithm has the same recognition rate in MIT and Yale database because the test faces have the same illumination pattern as those in the bootstrap set of quotient image. But the quotient image method loses its effects for Yale B Set 2 database and even can not recognize face images with small illumination variation, which are correctly recognized by direct correlation method. This result further proves that linear combination of three kinds of differently illuminated images as in equation (2) can not cope with images with different shadow patterns. In addition, any non-frontal illumination will cause shadows in face image, therefore the equation (1) is only valid in non-shadow face image, which makes the quotient image method unfit for dealing with large illumination variation.

Because the subjects in Yale B are in static state, the alignment is perfect, the recognition rate in Yale B set 1 is very high (comparing with results on Yale database and MIT database). These results also show that alignment will be a key point for this algorithm and they are consistent with the results of [4].

5. Conclusion

Our new algorithm takes advantages of quotient image's texture mapping ability and can synthesize more complex face images with large shadow than quotient image algorithm. Our algorithm needs only one face image associated with each person in the face database. The experiments demonstrate the algorithm can strikingly improve the recognition rate.

References

[1] Yael Adnin, Yael Moses and shimon Ullman, "Face recognition: The problem of compensating for changes in illumination direction", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, 1997.

[2] B. K . P. Horn, and M. Brooks, Seeing shape from shading, Cambridge, Mass, MIT press, 1989

[3] R. J. Woodham, " Photometric method for determining surface orientation from multiple images", Optical engineering", Vol. 19, No. 1, pp 139-144,1980

[4] Amnon Shashua, and Tammy Riklin-Raviv, "The quotient image: Class-based re-rendering and recognition with varying illuminations", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp129-139, 2001

[5] Athinodoros S. Georghiades and Peter N. Belhumeur, "From Few to many: Illumination cone models for face recognition under variable lighting and pose", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 6, pp 643-660, 2001

[6]M. Turk, A. Pentland, "Eigenface s for recognition", J. Cognitive Neuroscience, Vol. 3, pp. 71-86, 1991.

[7] R. Brunelli, "Estimation of pose and illuminant direction for face processing", Image and Vision Computing, volume 15, number 10, pp. 741-748, 1997

[8] Amnon Shashua, "Geometry and Photometry in 3D Visual recognition", PhD thesis, MIT, 1992.

[9] Georghiades, A.S. and Belhumeur, P.N. and Kriegman, D.J, "From Few to Many: Generative Models for Recognition under Variable Pose and Illumination", IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2000, pp277-284

[10] P. Belhumeur, D.J.Kriegman, and A.L.Yuille," The bas-relief ambiguity", IEEE Conf. on Computer Vision and Pattern Recognition, pages 1060--6, 1997