Illumination Invariant Algorithm for Face Image Based on Tricolor Attenuation Theory

Dawei Yang, Siyuan He, Jiandong Tian, Yandong Tang

Abstract—In the paper, a tricolor attenuation theory is introduced, a new algorithm which transform a color face image into a gray one which is illumination invariant based on the tricolor attenuation theory is proposed. The color face image is segmented into shadow part and non shadow part using tricolor attenuation theory, then computing the change rate between shadow region and non shadow region in each color channel, and structure a operator T with the changed rate. Finally, applying the operator T to the whole face image. The experiments demonstrate the validation of the algorithm.

I. INTRODUCTION

Face detection and face recognition are important in many applications such as security, human-machine interfaces, gesture based computer interfaces, automatic driver monitoring, biometric identification and video search. Face detection is an essential component in human-computer interaction, video surveillance, face tracking and face recognition[1-3]. Face images are significantly changed by lighting conditions and therefore may cause performance degradation both in face detection and face recognition[4]. Illumination cannot be controlled and varies according to the daylight conditions in many real applications. Moreover, strong shadows cast from a direct source can make facial features invisible.

Many methods have been proposed to solve the problems of the varying illumination of face images.

In the feature-based approaches, faces are represented by illumination invariant features. Typically, these are geometrical measurements and relationships between local facial features such as the eyes, mouth, nose and chin. This method uses the image information including the edge information, image convolved by 2D-Gabor filter, gray-level difference image and transformed image by Log operator[5]. These methods change the gray level contrast and texture information of image. The feature-based methods are known to be robust against varying illumination conditions. However, they rely on the accurate face and facial feature detection. The illumination invariant features are limited and not enough for recognition in large scale face databases. Under the largely varied lighting conditions, these methods achieve low level performance.

In holistic methods, the entire face image is considered as face representation without taking into account any specific geometrical features. A face image could be thought of as a point in a high dimensional image space. To avoid computational complexities and to reduce redundant data, face images are first linearly transformed into a low dimensional subspace before extracting a feature vector. The most commonly used dimension reduction technique is the Principal Component Analysis (PCA). PCA leaves out the first 3 eigenfaces could reduce the effect of variations in illumination [5]. However, it may also lead to the loss of the information useful for accurate identification, and these methods need a large number of samples to learn the varied lighting spaces, therefore, there should be enough training images.

The generative methods have been utilized to address the problem of varying illumination conditions in face recognition based on the assumption of the Lambertian model. Belhumeur and Kriegman[6] demonstrate that a set of images of an object under fixed posed, consisting of diffuse reflection components and shadows under arbitrary lighting conditions forms a convex cone, termed the illumination cone, in the image space and that this illumination cone can be approximated by a low-dimensional subspace. These generative methods perform well under varying illumination conditions, but they require a number of training samples which represent extreme illumination conditions. It may be possible to acquire a number of training images in the certain applications such as ID cards and drivers license, but difficult to get those in the surveillance and the counter terrorism related applications.

Now the most commonly used pre-processing method is generating an illumination invariant image and using this image to assist the face recognition. The methods based on Retinex theory[7] have been applied to the illumination compensation. The theoretic foundation of the Retinex theory is that an image I is regarded as the product I = L * R, where R is the reflectance and L is the illumination. The nature of L is determined by the illumination source, whereas R is determined by the characteristics of the imaged objects. Therefore, the illumination normalized images for face recognition can be achieved by estimating the illumination L and then dividing the image I by it. In the most Retinex methods, the reflectance R is estimated as the ratio of the image I and its smooth version which serves as...
the estimate of the illumination $L$ [8]. These Retinex methods have the common advantages: They don’t require training images and have relatively low computational complexity. However, these methods still cannot completely remove the cast shadows.

The Quotient Image (QI) method [9] uses a training set of face samples to generate a linear combination for each pixel. The ratio of albedo between the original face image and the linear combination is illumination invariant. QI method could improve the performance of face recognition, but it needs a large training set of face images and this is not practical under many conditions. The Self-Quotient Image model (SQI) [10] is based on the basic conception of QI. SQI does not need a linear combination of face samples. It uses the low-frequency part which represents the illumination effects of face image to divide the original image, and then generate the illumination invariant features. The SQI method neither uses the information about the lighting source, nor needs a training set. It directly extracts the illumination invariant face features. But the edge information is over-smoothed by the smoothing filter; therefore, the image processed by SQI loses some features.

In this paper, we introduce a tricolor attenuation theory. Based on the theory, we propose an algorithm to transform a color face image to a gray one which is not sensitive to shadows so much compared with the original one. The paper is organized as follows. In Section II, we introduce the tricolor attenuation theory. In Section III, specify the algorithm based on the tricolor attenuation theory. In Section IV, the experiments on the facade images taken outdoors are done, followed by the analysis of the experimental results. Section V is the conclusion.

II. TRICOLOR ATTENUATION THEORY

Light source is the key element for imaging. In outdoor scenes, there are mainly two light sources: direct sunlight, which can be regarded as a point light source; diffuse skylight: which can be regarded as an area light source. Shadows will occur when direct light from a light source is partially or totally occluded. Shadow can be divided into two types: self shadow and cast shadow. The self shadow is the part of an object that is not illuminated by direct light; the cast shadow is the dark area projected by an object on the background. Cast shadow can be further divided into umbra and penumbra region. The umbra region is the part of a cast shadow where direct light is completely blocked; the penumbra region is the part of a cast shadow where direct light is only partially blocked. As shown in Fig. 1, the illumination on non-shadow region is daylight (direct sunlight and diffused skylight); that on penumbra is skylight and part of sunlight; that on umbra is only skylight. Since skylight is a component of daylight, pixel intensity in shadow must be lower than that in non-shadow background[11].

Denoting $[F_R \ F_G \ F_B]$ as a tricolor vector of a pixel value in a color image $F$, $[F_{NS_R} \ F_{NS_G} \ F_{NS_B}]$ as a pixel value vector in a non-shadow background region, $[F_{S_R} \ F_{S_G} \ F_{S_B}]$ as a pixel value vector in the corresponding shadow region which has the same response of reflectance as $[F_{NS_R} \ F_{NS_G} \ F_{NS_B}]$, and $[\Delta R \ \Delta G \ \Delta B]$ as the value attenuation vector, the relationship between $[F_{NS_R} \ F_{NS_G} \ F_{NS_B}]$ and $[F_{S_R} \ F_{S_G} \ F_{S_B}]$ is,

\begin{align*}
\Delta R &= F_{NS_R} - F_{S_R} \\
\Delta G &= F_{NS_G} - F_{S_G} \\
\Delta B &= F_{NS_B} - F_{S_B}
\end{align*}

(1)

Formula (1) implies that if $\Delta R, \Delta G, \Delta B$ are different, the disparities of $R, G, B$ channels of a shadow region are expected to be different from those of the corresponding non-shadow background region. Taking $\Delta R > \Delta G > \Delta B$ as an example, if we subtract $B$ channel from $R$ channel,

\begin{align*}
F_{S_R} - F_{S_B} &= F_{NS_R} - \Delta R - (F_{NS_B} - \Delta B) \\
&= F_{NS_R} - F_{NS_B} + (\Delta B - \Delta R) < F_{NS_R} - F_{NS_B}
\end{align*}

(2)

Obviously, in this example, the disparity between $R$ and $B$ channels of shadow is lower than that of the corresponding non-shadow background.

The tricolor attenuation theory is that if we subtract the minimum attenuated channel from the maximum attenuated channel, the results in shadow regions will be lower than the results in non-shadow regions. This is very useful for shadow identification. In fact, in most cases, $\Delta R, \Delta G, \Delta B$ are different. The key problem is how to find the maximum and the minimum attenuated channels.

\[\max[R_S \ G_S \ B_S] < \max[R_{NS} \ G_{NS} \ B_{NS}]\] (3)

Formula (2) represents the tricolor attenuation theory, which can be used to segment a face image into a shadow part and a non-shadow part roughly.
III. ILLUMINATION INVARIANT ALGORITHM

Based on the segmentation of shadow and non-shadow of a face image, a new algorithm is proposed to transform a color face image into a gray face image without shadow.

Fig. 2. A shadow on background, in which \([r g b]\) is shadow region; \([R G B]\) is the corresponding non-shadow background.

As shown in Fig. 2, assuming \([r g b]\) is the shadow we have segmented, and \([R G B]\) is the corresponding non-shadow background. Assuming \(\Delta B \neq 0\), formula (4) always holds.

\[
(R - r) + (G - g) = (B - b) \cdot \frac{(R - r) + (G - g)}{(B - b)}
\]

\[
\Rightarrow (R - r) + (G - g) = (B - b) \cdot \left(\frac{\Delta R}{\Delta B} + \frac{\Delta G}{\Delta B}\right)
\]

So,

\[
R + G - \left(\frac{\Delta R}{\Delta B} + \frac{\Delta G}{\Delta B}\right) \cdot B = r + g - \left(\frac{\Delta R}{\Delta B} + \frac{\Delta G}{\Delta B}\right) \cdot b
\]

Denoting \(T\) operator as \(T = \begin{bmatrix} 1 & 1 & -\left(\frac{\Delta R}{\Delta B} + \frac{\Delta G}{\Delta B}\right) \end{bmatrix}\), equation (5) can be rewritten as \(T[R G B] = T[r g b]\). We find that after \(T\) operation, the shadow pixels and the non-shadow pixels are strictly equal. Apparently, we obtain a shadow invariant image. The steps of the algorithm as follows:

Step 1: Computing the tricolor attenuation image. In a color image, when suppose \(R > B\), it is not sure that each pixel’s intensity of R channel is always larger than which of B channel, so we select the maximum and the minimum channel intensity of each pixel, then getting the tricolor attenuation image using formula (6).

\[
\]

Step 2: Segment the face image into two parts: one is with shadow and the other is without shadow using formula (7).

\[
g(x, y) = \begin{cases} 1 & G(x, y) > T_F \\ 0 & \text{others} \end{cases}
\]

Where \(G\) is the tricolor attenuation image gained in step 1, \(T_F\) is the threshold to segment the shadow and now shadow part, which value can be computed automatically.

Step 3: Compute operator \(T\). The key work of the step is to compute \([\Delta R \quad \Delta G \quad \Delta B]\). Selecting a narrow region around the shadow part segmented in step 2 as non-shadow region, then computing \(T\) using the average of shadow and non-shadow region.

Step 4: Applying operator \(T\) to the original color face image and getting a shadow insensitive face image \(F'\).

Step 5: Normalizing the \(F'\) into the range \([0, 1]\) using \(F_{\text{NORM}} = (F' - F'_{\text{min}})/(F'_{\text{max}} - F'_{\text{min}})\). Where \(F'_{\text{min}}\) and \(F'_{\text{max}}\) are the minimal and maximal values of \(F'\).

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed algorithm, we take some outdoor and indoor face images with a digital camera, and do some experiments. Fig. 3 shows the results of the proposed algorithm. The left is the original color face image with cast shadow in his face. The middle is the tricolor attenuation image. A threshold of the image is computed using Otsu’s method[12]. The right is the result of the proposed algorithm.

We also do some comparison with other commonly used illumination compensation method. Fig. 4 shows some compensation results between the proposed algorithm and SQI method. Fig. 4 (a) shows two color face images derived by a camera outdoors, and cast shadows on their faces. Fig. 4 (b) shows the results of proposed algorithm, which are gray images and shadows on which have vanished, while the facial features such as eyes, mouth, nose and part of contour of the face keep clearly. Fig. 4 (c) shows the result of SQI, shadows vanish and facial features remain, but part of the contour of face is fuzzy, and the contrast ratio of the whole image is lower than that of proposed.
Fig. 5 shows the result when applying the proposed algorithm to indoor face image. The color face image is getting in a darkroom and the light source is lamplight. The result indicates that the proposed algorithm is not appropriate for the indoor face image, the condition of the image is not match with the condition of the tricolor attenuation theory.

V. CONCLUSION

The shadows on a face image will affect the detection and recognition of face image seriously. In this paper, we introduce a tricolor attenuation theory, and propose a new algorithm to transform a color face image to a gray one which is illuminant invariant based on the tricolor theory. The experiments of face images outdoors demonstrate the availability of the algorithm.

The algorithm make a good use of color channel information, it is valid for color face image outdoors, while not fit for indoor face image. The future work is to study the algorithm and make it suited to indoor face images.

REFERENCES


