A NEW METHODOLOGY OF ILLUMINATION ESTIMATION/NORMALIZATION BASED ON ADAPTIVE SMOOTHING FOR ROBUST FACE RECOGNITION

Young Kyung Park, Joong Kyu Kim

School of Information and Communication Engineering, SungKyunKwan University.
300, Chun-Chun-Dong, Chang-An-Ku, Suwon, Korea

ABSTRACT

In this paper, we propose a novel method of illumination estimation /normalization based on adaptive smoothing, which is to be applied to robust face recognition. In order to estimate the illumination in the framework of retinex theory, adaptive smoothing is applied based on both iterative convolution and two discontinuity measures. In addition to that, we also introduce a couple of new concepts, which are designed to be suitable especially for face images. One is the new conduction function for adaptive weighting, and the other is the smoothing constraint for more accurate description of real environments. The evaluations, which are conducted based on the Yale face database B, show that the proposed method achieves high recognition rates even in more challenging environments such as the case of using images with the worst case of illumination as a training set.

Index Terms— Adaptive smoothing, Illumination Estimation, Illumination Normalization, Retinex

1. INTRODUCTION

Illumination normalization is a major requirement in the face recognition process. In recent years, several methods which are based on retinex theory have been proposed because they have common advantages that they only require one training image and has relatively low computational complexity[1][2].

The retinex theory is based on the physical imaging model, in which an image $I(x, y)$ is regarded as the product of the reflectance $R(x, y)$ and the illumination $L(x, y)$ at each pixel (x, y)[1]. Therefore, the illumination normalization can be achieved by estimating the illumination L and then dividing the input image I by it. However, it is impossible to estimate L from I, unless something else is known about either L or R. Hence, various assumptions and simplifications about L, or R, or both are proposed to solve this problem[1]. A common assumption is that edges in the scene are edges in the reflectance, while illumination spatially changes slowly in the scene. Thus, in most retinex based algorithms, L is estimated as the smooth version of I. Single Scale Retinex (SSR), the latest version of Land's retinex, employed a simple Gaussian filter to estimate L[3]. However, strong cast shadows violate the assumption of slowly varying illumination, and halo effects are often visible at large illumination discontinuities in I. To solve this problem, Jobson et al. reduced the halo effects by combining several low-pass filtered copies for the estimation of L[4]. Discontinuity preserving filter is also one of the most efficient methods that can estimate L. For example, self-quotient image(SQI) improved performance for illumination problem using weighed Gaussian filter[2]. However, these methods still cannot completely remove large illumination discontinuities and ultimately cannot avoid a decrease of the recognition rate. Therefore, more efficient discontinuity preserving filter must be employed to estimate L.

In this paper, we propose a new method to estimate and normalize illumination applying recently developed concepts of discontinuity preserving filters in the framework of retinex based method. Our method is mainly based on adaptive smoothing using iterative convolution employing the idea that combines two discontinuity measures[5][6]. For estimation of illumination especially suitable to face recognition, we introduce two new useful concepts. First, we propose a new form of conduction function that is used to determine a discontinuity level in each pixel (x, y). Then, we apply additional constraint for more accurate description of real environments.

The paper is organized as follows. In section 2, the proposed method is described in detail. In section 3, the experimental results are presented. Finally a conclusion is made in section 4.

2. ILLUMINATION ESTIMATION/NORMALIZATION BASED ON ADAPTIVE SMOOTHING

The key idea of adaptive smoothing is to iteratively convolve the input image to be smoothed with a 3×3 averaging mask whose coefficients reflect, at each point, the discontinuity level of the input image. As verified in [6], this adaptive smoothing process is the approximation of anisotropic diffusion. General framework of adaptive smoothing can be formulated as follows. Since we estimate illumination L as smoothed version of input image I, the initial value of the estimated illumination, i.e. $L^{(0)}(x, y)$, must be the same as $I(x, y)$. Then, the estimated illumination $L^{(t+1)}(x, y)$ at the $(t + 1)th$ iteration

is given by:

$$L^{(t+1)}(x,y) = \frac{1}{N^{(t)}} \sum_{i=-1}^{1} \sum_{j=-1}^{1} L^{(t)}(x+i,y+j)w^{(t)}(x+i,y+j)$$

(1)

with

$$N^{(t)} = \sum_{i=-1}^{1} \sum_{j=-1}^{1} w^{(t)}(x+i,y+j)$$

(2)

where

$$w^{(t)}(x,y) = g(d^{(t)}(x,y))$$

(3)

$N^{(t)}$ in (2) represents a normalizing factor. A conduction function g is a nonnegative monotonically decreasing function such that $g(0) = 1$ and $g(d^{(t)}(x,y)) \to 0$ as $d^{(t)}(x,y)$ increases, where $d^{(t)}(x,y)$ represents the amount of discontinuity at each pixel (x,y). For more efficient setting of $w^{(t)}(x,y)$, unlike in (3), we employ the idea that combines two discontinuity measures[5].

2.1. Discontinuity Measures

To determine the level of discontinuity at each pixel (x,y), two measurements of discontinuities are used[5]: spatial gradient and local inhomogeneity.

2.1.1. Spatial Gradient

Spatial gradient is a common local discontinuity measure in image processing. The spatial gradient of an image $I(x,y)$ at pixel (x,y) is defined as the first partial derivatives of its image intensity function with respect to x and y:

$$\nabla I(x,y) = [G_x, G_y] = [\frac{\partial I(x,y)}{\partial x}, \frac{\partial I(x,y)}{\partial y}]$$

(4)

where the derivatives are approximated by

$$G_x = I(x+1,y) - I(x-1,y)$$

(5)

$$G_y = I(x,y+1) - I(x,y-1)$$

(6)

The magnitude of the gradient vector in (4), on the other hand, is given by

$$|\nabla I(x,y)| = \sqrt{G_x^2 + G_y^2}$$

(7)

2.1.2. Local Inhomogeneity

In addition to spatial gradient, Chen proposed to use inhomogeneity as another measure of discontinuity[5]. Although this measure is very efficient, it is very time consuming. Therefore, we propose simpler measure for real time application. Another measure that we propose is just the average of local intensity differences for each pixel (x,y) in the face image. We call this measure local inhomogeneity. This measure provides the degree of uniformity for all the pixels in the
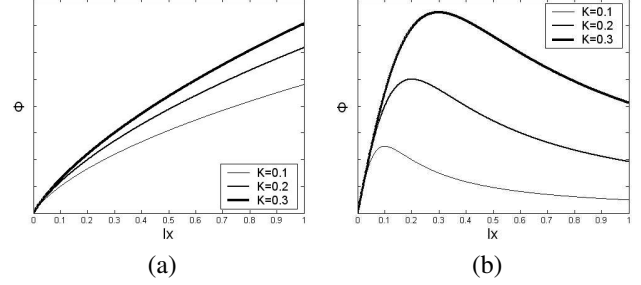

Fig. 1. (a)Flow magnitude functions by (15). (b)Flow magnitude functions by (14).

small neighborhood of current pixel. If local inhomogeneity at pixel(x,y) is large, we can expect discontinuity occurs at pixel(x,y). The average of local intensity differences at pixel(x,y) is

$$\tau(x,y) = \frac{\sum\sum_{(m,n)\in\Omega} |I(x,y) - I(m,n)|}{|\Omega|}$$

(8)

where Ω is a convolution region, and (m,n) indicates the locations of the neighboring pixels. Then, the average $\tau(x,y)$ at each pixel(x,y) is normalized by

$$\hat{\tau}(x,y) = \frac{\tau(x,y) - \tau_{min}}{\tau_{max} - \tau_{min}}$$

(9)

where τ_{max} and τ_{min} respectively are the maximal and minimal values of τ across the entire face image. To emphasize higher values that more likely correspond to cast shadow, we also adopt a nonlinear transformation as follows.

$$\tilde{\tau}(x,y) = \sin(\frac{\pi}{2}\hat{\tau}(x,y)), 0 \le \tilde{\tau}(x,y) \le 1$$

(10)

Although this measure is calculated based only on its nearest neighborhood, we can expect the propagation effect of local inhomogeneity incurred by iterative convolution.

2.2. Conduction Function and Smoothing Constraint

To utilize two discontinuity measures described in the previous section, the proper conduction function g must be defined. As mentioned, g is nonnegative monotonically decreasing function because a large weight should be assigned to a pixel that involves low discontinuity, and vice versa for discontinuity preserving. We apply nonnegative monotonically decreasing function g to both spatial gradient and local inhomogeneity as follows.

$$\alpha(x,y) = g(\tilde{\tau}(x,y), h)$$

(11)

$$\beta(x,y) = g(|\nabla I(x,y)|, S)$$

(12)

Here, $h(0 < h < 1)$ and $S(S > 0)$ are used to determine level of discontinuities which must be preserved. Among possible selections of g, most works[5][6] proposed to use two forms, given below in (13) and (14).
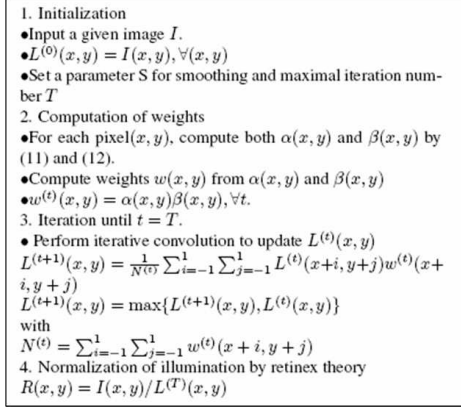
```
1. Initialization
•Input a given image I.
•L^(0)(x,y) = I(x,y), ∀(x,y)
•Set a parameter S for smoothing and maximal iteration num-
ber T
2. Computation of weights
•For each pixel(x,y), compute both α(x,y) and β(x,y) by
(11) and (12).
•Compute weights w(x,y) from α(x,y) and β(x,y)
•w^(t)(x,y) = α(x,y)β(x,y), ∀t.
3. Iteration until t = T.
• Perform iterative convolution to update L^(t)(x,y)
L^(t+1)(x,y) = (1/N^(t)) Σ^1_{i=-1} Σ^1_{j=-1} L^(t)(x+i,y+j)w^(t)(x+
i,y+j)
L^(t+1)(x,y) = max{L^(t+1)(x,y), L^(t)(x,y)}
with
N^(t) = Σ^1_{i=-1} Σ^1_{j=-1} w^(t)(x+i,y+j)
4. Normalization of illumination by retinex theory
R(x,y) = I(x,y)/L^(T)(x,y)
```

Fig. 2. Proposed illumination estimation/normalization

$$g(z, K) = \exp(-\frac{z^2}{2K^2}) \tag{13}$$

$$g(z, K) = \frac{1}{1 + (\frac{z}{K})^2} \tag{14}$$

These two functions cause two different operations in adaptive smoothing as iteration proceeds: one is the smoothing within homogeneous regions, and the other is the sharpening of strong discontinuities(e.g. edges) that will be preserved. However, edge sharpening operation is not proper to our objective although this is very efficient in the view of feature extraction(e.g. edge detection). Because edge sharpening operation makes boundary of cast shadow more sharp, the boundary between shadow region and non-shadow region may be more visible after normalization. Therefore, we propose a new form of g without edge sharpening effect.

$$g(z, K) = \frac{1}{1 + \sqrt{\frac{z}{K}}} \tag{15}$$

In order to see that this function does not have the edge sharpening effect, let us recall that the adaptive smoothing is the approximation of the anisotropic diffusion and consider only 1-D case without loss of generality. Then, smoothing process can be formulated as follows.

$$\frac{dI(x,t)}{dt} = \frac{d}{dx}(\phi(I_x(x,t)) \tag{16}$$

with

$$\phi(I_x(x,t)) = g(I_x(x,t), K)I_x(x,t) \tag{17}$$

Perona and Malik described qualitatively the shape of the flow magnitude function ϕ that causes edge sharpening[7]. As shown in figure 1, the flow magnitude function with the existing conduction function((14)) has finite optimum that ϕ is monotonically increasing for $|I_x| < K$ while it is monotonically decreasing for $|I_x| > K$. However, the flow magnitude function with proposed conduction function((15)) is a throughout-increasing function and possesses no finite optimum. As described in [7], the edge sharpening effect happens only when ϕ is monotonically decreasing. Therefore,
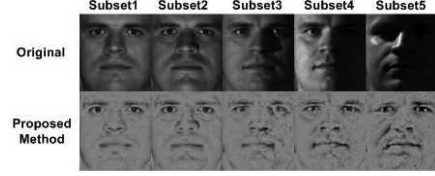

Fig. 3. Examples of face images normalized by the proposed method.

our proposed conduction function g has not the edge sharpening effect. This means that even strong discontinuities will be smoothed eventually. However, we can expect the preservation of strong discontinuities because iteration number, which is required for illumination normalization of a face image, is very small(10 in our experiments).

The values of S and h can be selected by many experiments, and we give the proper values of $S(S = 1)$ and $h(h = 0.1)$ for face image normalization. Now, the corresponding weights of the convolution mask $w^{(t)}(x, y)$ are determined using $\alpha(x, y)$ and $\beta(x, y)$ as follows:

$$w^{(t)}(x, y) = \alpha(x,y)\beta(x,y), \forall t \tag{18}$$

For more accurate description of real environments, we address an additional constraint that surfaces cannot reflect more light than what is shed on them. We call this constraint smoothing constraint. Thus, the reflectance R must be always smaller than unity. Figure 2 summarizes the steps of our proposed method for estimating and normalizing illumination, where both smoothing constraint and adaptive weighting are applied.

3. EXPERIMENTAL RESULTS

In order to evaluate the robustness and effectiveness of the proposed method, we used the images from Yale face database B[8], and computed the recognition accuracies using PCA algorithm[9]. The proposed method is also compared with three other existing methods of illumination normalization: SSR[3], SQI[2] and histogram equalization.

The Yale face database B contains 5,760 images taken from 10 subjects under 576 viewing conditions(9 poses × 64 illumination conditions). Since we are only concerned with the illumination problem in this paper, we selected 640 images for 10 subjects representing 64 illumination conditions under the frontal pose. Images in the database are divided into 5 subsets based on the angle of the light source directions. The 5 subsets are subset 1($0°$ to $12°$), subset 2($13°$ to $25°$), subset 3($26°$ to $50°$), subset 4($51°$ to $77°$), and subset 5(above $78°$). We refer to images in both subset 4 and subset 5 as images with large illumination variation. Figure 3 shows examples of face images normalized by the proposed method.

First, we used subset 1 as training set and tested other subsets. As shown in figure 4(a), the proposed method achieved
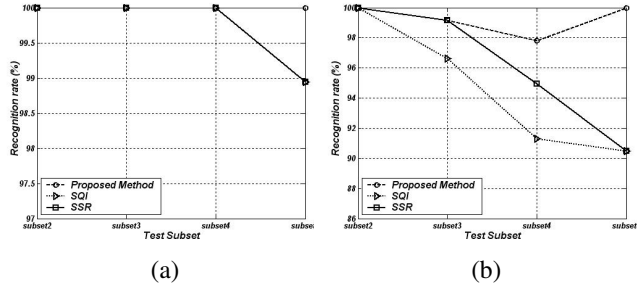
Fig. 4. Recognition accuracies(%) using (a)subset 1, and (b)only the ideal images, as the training set.
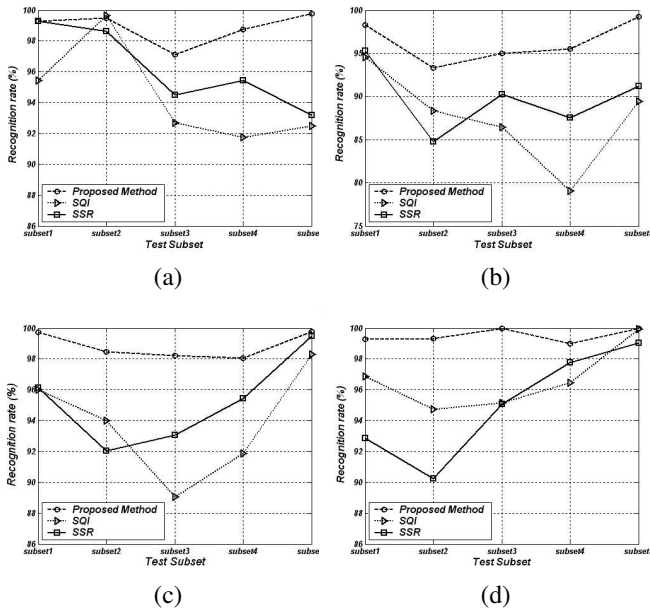

Fig. 5. Recognition accuracies(%) using (a)subset 2 ($test2$), (b)subset 3 ($test3$), (c)subset 4 ($test4$), and (d)subset 5 ($test5$) as the training set.

recognition rates of 100% in all subsets. Then, we only used the 10 ideal images (one($0°$) for each subject in subset 1) as a training set and tested other subsets. The results are given in figure 4(b), and it is clear that the proposed method outperforms other methods. Finally, we used images in each of the other subsets (2 to 5) as training set. Since subset 2 to 5 represent the illumination conditions close to real environments, we can say that this test is more meaningful and practical than two previous tests. As shown in figure 5, four tests are carried out. Each $testX$ denotes that the training set is from subset X, where $X = 2, 3, 4, 5$, and the test is done for all subsets. In each $testX$, 10 trials with other training set are averaged, and only 10 images(one for each subject) are used for training in each trial. From the figure 5, it is clear that the proposed method has consistent and promising results even when images with large illumination variation (subset 4 and subset 5) are used as the training set. Our method achieved high average recognition rate of 98.52% from all tests which only used 10 images(one for each subject) as training set.

4. CONCLUSION

In this paper, we proposed a novel method to estimate and normalize illumination for robust face recognition. The essence of our proposed method is the illumination estimation using a new adaptive smoothing which is designed specially for face recognition. For this purpose, we introduced couple of new concepts; new conduction function and the additional constraint that are decided to be suitable to face images. Using the proposed method, we showed that even images with strong shadow are effectively normalized, and consequently some remarkable improvements in recognition capability are observed and compared to the existing methods. We can apply the proposed method to any single face image without any prior information.

5. REFERENCES

[1] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," *4th AVBPA*, pp. 10–18, 2003.

[2] H. Wang, S. Li, and Y. Wang, "Generalized quotient image," *IEEE CVPR*, 2004.

[3] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. on Image Processing*, vol. 6, pp. 451–452, 1997.

[4] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. on Image Processing*, vol. 6, pp. 965–976, 1997.

[5] K. Chen, "Adaptive smoothing via contextual and local discontinuities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27(10), pp. 1552–1567, 2005.

[6] P. Saint-Marc, J. S. Chen, and G. Medioni, "Adaptive smoothing: a general tool for early vision," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13(6), pp. 514–529, 1991.

[7] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12(7), pp. 629–639, 1990.

[8] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23(6), pp. 643–660, 2001.

[9] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3(1), pp. 71–86, 1991.