Available online at www.sciencedirect.com
ScienceDirect
Signal Processing 88 (2008) 1929–1945
SIGNAL PROCESSING
www.elsevier.com/locate/sigpro

Retinex method based on adaptive smoothing for illumination invariant face recognition

Young Kyung Park, Seok Lai Park, Joong Kyu Kim*

School of Information and Communication Engineering, Sungkyunkwan University, 300, Chun-Chun-Dong, Chang-An-Ku, Suwon 440-746, Republic of Korea

Received 7 August 2007; received in revised form 25 January 2008; accepted 25 January 2008
Available online 15 February 2008

Abstract

In this paper, we propose the Retinex method for illumination invariant face recognition developed on the basis of adaptive smoothing technology. By the well-known Retinex theory, illumination is generally estimated and normalized by smoothing the input image first and then dividing the estimate into the original input image. Therefore, performance mainly depends on how good the estimated illumination is. The proposed method estimates illumination by iteratively convolving the input image with a 3×3 smoothing mask weighted by a coefficient via combining two measures of the illumination discontinuity at each pixel. We address a couple of additional concepts, which are designed to be suitable especially for face images. One is the new conduction function for adaptive smoothing, and the other is the smoothing constraint for more accurate description of real environments. In this way, we can achieve an efficient illumination normalization in which face images with even strong shadows are normalized efficiently. The proposed method is evaluated based on Yale face database B, CMU PIE database and AR face database by applying PCA. The comparative results indicate that the proposed method present consistent and promising results even when images under harsh illumination conditions are used as a training set.

Keywords: Face recognition; Illumination normalization; Adaptive smoothing; Iterative convolution; Retinex; Anisotropic diffusion; Conduction function; Smoothing constraint

1. Introduction

Over the last decade, significant progress has been achieved in face recognition area. Principle component analysis (PCA) [1], linear discriminant analysis (LDA) [2], discriminative common vector (DCV) [3], etc. have been developed successfully for face recognition, the product of which has mostly been commercialized. However, for reliable face recognition in unconstrained environment, there still remain many tasks to be solved in the area. Among many factors affecting the performance of face recognition systems, illumination is known to be the one of the most significant. For example, ambient lighting varies greatly during the course of the day, and from one day to another, as well as between indoor and outdoor environments. Moreover, strong shadows cast from a direct light source can make certain facial features invisible. Therefore,

*Corresponding author. Tel.: +82 31 290 7122; fax: +82 31 290 7687.
E-mail addresses: multipym@skku.edu (Y.K. Park), jkkim@skku.edu (J.K. Kim).

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2008.01.028

illumination normalization is a major requirement in the face recognition process and also is a central topic in the field of computer vision.

In recent years, numerous approaches have been proposed to solve illumination problems in face recognition. Georghiades et al. [4] showed that the illumination cones of human faces can be approximated well by low dimensional linear subspaces. Therefore, the set of face images in fixed pose but under different illumination conditions can be efficiently represented using an illumination cone. Recently, Basri and Jacobs [5] showed that the set of images of a convex Lambertian object obtained under a variety of lighting conditions can be well approximated using a 9D linear subspace which is formed by nine harmonic images. However, this method requires knowledge of the object's surface normals and albedos before the harmonic subspace can be computed. As nine images are not real images, Lee et al. [6] showed that the linear subspaces could be directly generated using real nine images captured under a particular set of illumination conditions. Based on these discoveries, some relighting methods have also been proposed to generate photo-realistic images under new lighting conditions [7,8]. These methods provide the new insight to facilitate face recognition under different lighting conditions. However, the above methods still either require certain assumptions to be made about the light source or need a large number of training sets, and these requirements are not considered practical in real applications. On the other hand, there are alternative methods available which are based on Retinex theory although these methods are not designed for face recognition initially. The Retinex theory motivated by Land [9] is based on the physical imaging model, in which an image $I(x,y)$ is regarded as the product $I(x,y) = R(x,y)L(x,y)$ where $R(x,y)$ is the reflectance and $L(x,y)$ is the illumination at each pixel (x,y). Here, the nature of $L(x,y)$ is determined by the illumination source, whereas $R(x,y)$ is determined by the characteristics of the imaged objects. Therefore, the illumination normalization for face recognition can be achieved by estimating the illumination L and then dividing the 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 [10]. 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 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, and many smoothing filters to estimate the illumination have been proposed for illumination invariant face recognition. Single scale Retinex (SSR), the latest version of Land's Retinex that was implemented and tested by Jobson et al. [11], employed a simple linear filter with Gaussian kernel. However, strong shadows cast from a direct light source violate the assumption that the illumination slowly varies, and halo effects are often visible at large illumination discontinuities in I. To solve this problem, Jobson extended SSR to multiscale Retinex (MSR) [12] by combining several low-pass filtered copies of the logarithm of I using different cut-off frequencies for each low-pass filter. Recently, Gross and Brajovie [10] introduced an anisotropic filter to reduce these halo effects to some extent. More recently, self-quotient image (SQI) [13,14] has been proposed with impressive improvement of performance for illumination problem. SQI employs the weighed Gaussian filter in which the convolution region is divided into two sub-regions with respect to a threshold, and separate values of weights are applied in each sub-region. These Retinex methods have common advantages that they do not require training images and has relatively low computational complexity. However, these methods still cannot completely remove cast shadows because they lack in adaptability which can preserve discontinuities efficiently. As a result, they ultimately cannot avoid the decrease in the recognition rate. To solve these problems, efficient discontinuity preserving filter must be employed to estimate L. Though not developed for face recognition, many discontinuity preserving filters such as adaptive smoothing [15,16], low curvature image simplifier (LCIS) [17], bilateral filtering [18] have been proposed.

In this paper, we propose a novel Retinex method for illumination invariant face recognition applying recently developed concepts of discontinuity preserving filters. Our method is mainly based on adaptive smoothing using iterative convolution employing the idea that combines two discontinuity measures. For estimation of illumination especially suitable to face recognition, we introduce two additional concepts: first, we propose a new form of conduction function that is used to determine a discontinuity level in each pixel (x,y), and then we apply