Neurocomputing 199 (2016) 16–30

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Face recognition under varying illuminations using logarithmic fractal dimension-based complete eight local directional patterns

Mohammad Reza Faraji*, Xiaojun Qi

Department of Computer Science, Utah State University, Logan, UT 84322-4205, USA

A R T I C L E I N F O

Article history:
Received 8 October 2015
Received in revised form
6 January 2016
Accepted 31 January 2016
Available online 16 March 2016

Keywords:
Face recognition
Illumination variations
Complete eight local directional patterns
Logarithmic fractal dimension

A B S T R A C T

Face recognition under illumination is really challenging. This paper proposes an effective method to produce illumination-invariant features for images with various levels of illumination. The proposed method seamlessly combines adaptive homomorphic filtering, simplified logarithmic fractal dimension, and complete eight local directional patterns to produce illumination-invariant representations. Our extensive experiments show that the proposed method outperforms two of its variant methods and nine state-of-the-art methods, and achieves the overall face recognition accuracy of 99.47%, 94.55%, 99.53%, and 86.63% on Yale B, extended Yale B, CMU-PIE, and AR face databases, respectively, when using one image per subject for training. It also outperforms the compared methods on the Honda UCSD video database using five images per subject for training and considering all necessary steps including face detection, landmark localization, face normalization, and face matching to recognize faces. Our evaluations using receiver operating characteristic (ROC) curves also verify the proposed method has the best verification and discrimination ability compared with other peer methods.

1. Introduction

Face recognition with a wide range of applications in security, forensic investigation, and law enforcement is a good compromise between reliability and social acceptance [1,2]. A typical face recognition system includes four main stages (shown in Fig. 1 [1]): face detection, landmark localization, face normalization, and face matching. In the first stage, face detectors such as the SNoW-based (sparse network of windows) face detector [3], the Viola-Jones detector [4], and the skin-color-pixel-based Viola-Jones face detector [5] detect faces in images or video frames. In the second stage, respective facial landmarks are localized for each detected face. Some of recently proposed facial landmark localization techniques include boosted regression with Markov networks (*BoRMaN*) [6], local evidence aggregation for regression-based facial point detection (*LEAR*) [7], discriminative response map fitting (*DRMF*) [8], and *Chehra v.1* (meaning "face" in hindi) [9]. In the third stage, detected faces are geometrically normalized and corrected by facial landmarks. First, *geometric rectification* is applied to normalize the pose. Pose normalization then uses facial landmarks to estimate the frontal view for each face [10]. Face images can be automatically cropped using facial landmarks, since landmarks exactly indicate the locations of key facial points in the image. The *illumination preprocessing* method is finally applied on the cropped images to obtain illumination-insensitive face images. In this stage, if face images contain facial expressions such as smile, anger, or scream, Gabor features-based expression normalization can be employed to extract facial features that are robust to expression variations [11,12]. In the final stage, a *face matching* process is used on the normalized faces to recognize the identity of the faces. This stage includes *feature extraction* and *face classification*. In *feature extraction*, each face is converted to a one-dimensional vector to be used for classification. Each face might also be divided into a regular grid of cells to extract histograms [13]. Dimension reduction methods can also be used to reduce the dimension of the data before a classifier is employed to verify the identity of the input face image. Representative dimension reduction methods include principal component analysis (PCA) [14–16], independent component analysis (ICA) [17,15], kernel principal component analysis (KPCA) [18], Fisher's linear discriminant analysis (FDA) [19], kernel Fisher's linear discriminant analysis (KFDA) [20,21], and singular value decomposition (SVD). Finally, the *face classification* stage uses the extracted features to classify and verify a new unknown face image with respect to the image database. Some of the popular classification methods used in face recognition systems are k nearest neighbors [22,23], neural-networks [24,25], and the sparse representation classifier (SRC) [26,27].

However, five key factors including illumination variations, pose changes, facial expressions, age variations, and occlusion significantly affect the performance of face recognition systems. Among these factors, illumination variations such as shadows,

* Corresponding author.
E-mail address: Mohammadereza.Faraji@aggiemail.usu.edu (M.R. Faraji).

http://dx.doi.org/10.1016/j.neucom.2016.01.094
0925-2312/

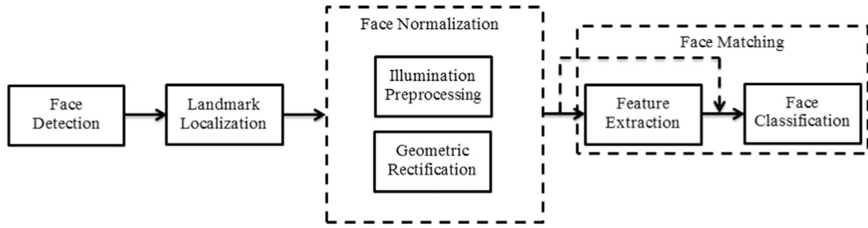

Fig. 1. Typical framework for illumination-invariant face recognition [1].

underexposure (too dark), and overexposure (too bright) attracted much attention in the last decade [28,1]. In the image of a familiar face, changing the direction of the light source leads to shifts in the location and shape of shadows, changes in highlights (i.e., making the face appear too bright or too dark), and reversal of contrast. Braje et al. [29] studied the effects of illumination in face recognition from the psychological perspectives. They found out that illumination variations have enormously complex effects on the image of an object. Specifically, varying the illumination direction resulted in larger image differences than did varying the identity of the face. Furthermore, cast shadows introduced spurious luminance edges that may be confused with object contours and consequently significantly hindered performance of face recognition systems. As a result, there is a need for extracting facial features which are invariant or insensitive to illumination variations. These invariant facial features can then be used for recognizing the face. Various methods have been proposed to deal with illumination variations. These methods can be categorized into the following three groups: gray-level transformation methods, gradient or edge extraction methods, and reflection field estimation methods [1,30].

Gray-level transformation methods redistribute the intensities in a face image with a linear or non-linear transformation function to correct the uneven illumination to some extent [1]. They are simple and computationally efficient and fast. But, they are not effective in eliminating the effects of illumination and therefore achieve an inferior performance than the methods in the other two groups. Representative methods include histogram equalization (HE) [31] and gamma intensity correction (GIC) [32]. HE adjusts and flattens the contrast of the image using the image's histogram. GIC maps the image $I(x,y)$ to the image $G(x,y)$ using $\alpha I(x,y)^{1/\gamma}$, where α is a gray-stretch parameter and γ is the Gamma coefficient.

Gradient or edge extraction methods produce illumination-insensitive representations by extracting gray-level gradients or edges from face images [1]. Representative methods include local binary patterns (LBP) [33], local directional patterns (LDP) [34], enhanced LDP (EnLDP) [35], local directional number patterns (LDN) [36], eight local directional patterns (ELDP) [37], adaptive homomorphic eight local directional patterns (AH-ELDP) [38], discriminant face descriptor (DFD) [39], local tetra patterns (LTrP) [40], enhanced center-symmetric local binary pattern (ECS-LBP) [41], and logarithmic fractal dimension (LFD) [30]. LBP takes P pixels in a circle of radius R around each pixel and threshold these pixels based on the value of the central pixel, where P and R are usually set to be 8 and 2 for face recognition applications, respectively. LDP, EnLDP, LDN, and ELDP produce eight directional edge images using Kirsch compass masks and encode the directional information to obtain noise and illumination-invariant representations. AH-ELDP filters face images using the homomorphic filter, enhances the filtered images using an interpolative function and then performs the ELDP method to produce final AH-ELDP images. The DFD method is an improvement of the LBP descriptor. It is involved with three-step feature extraction to maximize the appearance difference from different persons and minimize the difference from the same person. LTrP encodes the relationship among the center pixel and its neighborhood pixels using derivatives in vertical and horizontal directions. LFD performs a log function on face images and transfers images to the fractal dimension domain to produce illumination-invariant face representations.

Reflectance field estimation methods estimate the face reflectance field, which is illumination-invariant, from a face image. They usually apply the Lambertian reflectance-illumination model (i.e., retinex theory) [1,42] as their face imaging model. These methods are generally effective. However, the surface of the human face is approximately Lambertian, which makes illumination-invariant representation less sufficient [43]. Self-quotient image (SQI) [44,45], adaptive smoothing (AdaS) [46], Gradientface [47], Weberface [48], generalized Weberface (GWF) [49], and local Gradientface XOR and binary pattern (LGXBP) [50] are representative methods. SQI, which is a ratio image between a given test image and its smoothed version, implicitly indicates that each preprocessed image is illumination-invariant [51]. The AdaS method estimates the illumination by smoothing the input image using an iterative method. The Gradientface method creates a ratio image between the y-gradient and the x-gradient of a given image. The Weberface method, inspired by Weber's law and based on the Weber local descriptor (WLD) [52], creates a ratio image between the local intensity variation and the background. The Gradientface and Weberface methods are proven to be illumination-insensitive. The GWF method is a generalized multi-scale version of the Weberface method. It also considers an inner and outer ground for each pixel and assigns different weights for them to develop a weighted GWF (wGWF). The LGXBP method uses a histogram-based descriptor to produce illumination-invariant representation of a face image. It first transforms a face image into the logarithm domain and obtains a pair of illumination-insensitive components: gradient-face orientation and gradient-face magnitude. These two components are then encoded by LBP and local XOR patterns to form the local descriptor and the histogram for face recognition.

In this paper, we propose an effective method which encodes face images with various levels of illumination. To this end, we first filter images using the adaptive homomorphic filter to partially reduce the illumination. We then use the simplified logarithmic fractal dimension transformation as an edge enhancer technique to enhance facial features and remove illumination to some point. We finally produce eight edge directional images using Kirsch compass masks and propose a complete ELDP (CELDP), which considers both directions and magnitudes of the edge responses, to obtain illumination-invariant representations.

The contributions of the proposed method are: (1) Using the adaptive homomorphic filter to partially reduce the illumination by attenuating the low-frequency (i.e., illumination) component of each face image based on the characteristics of each face image. (2) Transforming filtered face images to the LFD domain by skipping every other adjacent scale (i.e., reducing the computational time by half) to enhance facial features such as eyes, eyebrows, nose, and mouth while keeping noise at a low level. (3) Employing a gradient-based descriptor which considers the relations among