Information Sciences 316 (2015) 234–245

Contents lists available at ScienceDirect

Information Sciences

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

Near infrared face recognition using Zernike moments and Hermite kernels

Sajad Farokhi [a,b], Usman Ullah Sheikh [a,*], Jan Flusser [b], Bo Yang [b]

[a] Digital Signal and Image Processing Research Group, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Johor, Malaysia
[b] Institute of Information Theory and Automation, Academy of Science of the Czech Republic, 18208 Prague 8, Czech Republic

ARTICLE INFO

Article history:
Received 21 June 2014
Received in revised form 7 April 2015
Accepted 11 April 2015
Available online 21 April 2015

Keywords:
Face recognition
Near infrared
Zernike moments
Hermite kernel
Decision fusion

ABSTRACT

This work proposes a novel face recognition method based on Zernike moments (ZMs) and Hermite kernels (HKs) to cope with variations in facial expression, changes in head pose and scale, occlusions due to wearing eyeglasses and the effects of time lapse. Near infrared images are used to tackle the impact of illumination changes on face recognition, and a combination of global and local features is utilized in the decision fusion step. In the global part, ZMs are used as a feature extractor and in the local part, the images are partitioned into multiple patches and filtered patch-wise with HKs. Finally, principal component analysis followed by linear discriminant analysis is applied to data vectors to generate salient features and decision fusion is applied on the feature vectors to properly combine both global and local features. Experimental results on CASIA NIR and PolyU NIR face databases clearly show that the proposed method achieves significantly higher face recognition accuracy compared with existing methods.

1. Introduction

With the rapid development of biometric technology, face recognition (FR) has become an active research area in the field of computer vision. Despite research efforts, FR remains a challenge principally because of large intra-class variations stemming from illumination conditions, facial expressions, wearing eyeglasses, head pose, scale and the effects of time lapse. In the last two decades, numerous techniques have been proposed to solve these challenges and to develop a more effective FR system [1,27,41,54,58,63,65,66]. Comprehensive reviews on recent FR methods were presented in [67,42]. Among the many issues in visible FR systems, variation of illumination is regarded as the most challenging for subject identification in cooperative, as well as non-cooperative, user scenarios [2,28,56,59]. Several illumination invariant FR methods have been proposed, including methods based on three-dimensional (3D) shapes of the face, as well as methods based on thermal images measuring body temperature [3,31,37,38,46,50]. However, 3D techniques are costly and require high computational complexity, while thermal images are extremely sensitive to environmental temperature, health conditions, perspiration and are opaque to eyeglasses [4,7,53]. Other methods to compensate for illumination problems have also been introduced, a typical example being presented in [52]. However, the solution to the illumination problem in FR is still not perfect. Recently, illumination variations have been addressed utilizing near infrared (NIR) imagery [10,24,34].

* Corresponding author. Tel.: +60 197055336.
E-mail addresses: fsajad2@live.utm.my (S. Farokhi), usman@fke.utm.my (U.U. Sheikh), flusser@utia.cas.cz (J. Flusser), yang@utia.cas.cz (B. Yang).

http://dx.doi.org/10.1016/j.ins.2015.04.030
0020-0255/

Several NIR FR methods have been proposed to achieve an accurate FR system. Li et al. based their advanced method on a local binary pattern (LBP) and statistical learning algorithms, achieving substantial improvement in their results [33]. However, their method suffered from the high sensitivity of LBP to noise and minor pose variations and alignment errors [9,13,25,35].

Zhang et al. utilized Gabor filtering to enhance image features, and subsequently extracted discriminative features using a directional binary code (DBC). DBC captures more spatial information than LBP, and thus gives greater recognition accuracy [64]. The authors were successful in proposing an accurate FR system. However, more challenges, such as wearing eyeglasses, still remain.

To improve recognition performance, both NIR and visible images and other types of modality, such as voice and fingerprints, have been attempted [57]. Shen et al. proposed a method based on boosted DBC features using both visible and NIR images [43]. DBC and the AdaBoost algorithm were used for feature extraction and classification, respectively, while the performance was improved by consequent information fusion on a decision level. Other similar approaches can be found in [20,51,55,61]. However, these remain beyond the scope of this paper.

Nighttime FR at long distance was recently investigated by Maeng et al. [36]. Their preprocessing step was based on the difference of Gaussians, then scale invariant feature transform (SIFT) and multi-scale local binary pattern (MLBP) were used for feature extraction. Their results indicated that SIFT has better accuracy compared to MLBP. Other improvements in this domain have also been presented in [26,36,40].

A broad review of infrared FR methods can be found in [18,19]. However, as reported by Ghiass et al., all papers are limited by failing to examine the performance of the proposed method in the simultaneous presence of all considered challenges. Most related works in the NIR domain have focused solely on the illumination problem, with scant attention paid to facial expression, wearing eyeglasses, variations in head pose and scale and the effects of time lapse. These factors are all known to introduce crucial problems in FR systems [6].

Farokhi et al. [12,13] systematically studied the possibility of NIR FR in the presence of the above mentioned challenges. In [12], they proposed a method based on Zernike moments (ZMs) to cope with face rotation and noise. Although their results were an improvement on those achieved by LBP [33], the authors concluded that an accurate system could not be based on single-type features. Instead, it should utilize a combination of both global and local features (fusion) to make the final decision. A similar conclusion was drawn in [15], also combining the global and local facial features presented in [32,62], and outperformed the individual use of these features. Farokhi et al. used this principle in NIR FR [13] in which ZMs were used to extract global features while undecimated discrete wavelet transform (UDWT) was used to extract local features. Their system achieved good recognition performance. However, UDWT is inefficient in terms of memory usage and computational time.

In this work, we follow this successful approach and replace UDWT by other local features that provide at least as good a discrimination power as UDWT while decreasing the computational time. We propose Hermite kernels (HKs) as filters to extract local features and ZMs as global features. Presented results of extensive experiments on the CASIA NIR and the PolyU NIR face databases show that our method achieves a higher recognition rate, compared to existing works in the presence of the most common challenges (face image modification). Our method is capable of overcoming and improving many of the shortcomings of the existing state-of-the-art techniques.

The paper is organized as follows: In Sections 2 and 3, brief reviews of ZMs and HKs are given. The proposed method is then described in Section 4. Experimental results and performance analysis are then presented in Section 5, and finally the paper is concluded in Section 6.

2. Zernike moments

ZMs were first introduced by Teague in the early 1980s, and have been applied in many research works [16,23,45,47]. ZMs belong to the family of so-called *radial moments*, whose basis functions are, in polar coordinates, products of a 1D polynomial in the radial direction and a harmonic function in the angular direction. As follows from the Fourier Shift Theorem, radial moments change under image rotation only in their phase, while the amplitude remains constant (the same is true also for quaternion moments, see [21]). This favorable behavior makes ZMs useful in coping with image rotations. On the contrary, the moments orthogonal on a rectangle, such as Legendre, Chebyshev, Gegenbauer and Krawtchouk [16,22], change under rotation in a much more complicated manner. This renders them quite difficult to utilize for constructing invariants. The ZMs have shown good performance in FR systems in the presence of facial expression, image rotation, and noise [11,12,14,30,44]. However, they cannot handle partial occlusions properly due to their global nature. The ZM of order p with repetition q of a function $f(r,\theta)$, where (r,θ) are polar coordinates, is defined by the following equation:

$$Z_{pq} = \frac{p+1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} V_{pq}^{*}(r,\theta) f(r,\theta) r dr d\theta, \quad |r| \leqslant 1, \tag{1}$$

where the symbol "$*$" denotes a complex conjugate and V_{pq} denotes a Zernike polynomial of order p and repetition q, which can be written as follows:

$$V_{pq}(r,\theta) = R_{pq}(r) e^{jq\theta}. \tag{2}$$