Information Fusion 22 (2015) 95–104

Contents lists available at ScienceDirect

Information Fusion

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

Face recognition based on pixel-level and feature-level fusion of the top-level's wavelet sub-bands

Zheng-Hai Huang [a,b], Wen-Juan Li [a], Jun Wang [a,c], Ting Zhang [a,b,*]

[a] Center for Applied Mathematics of Tianjin University, Tianjin 300072, PR China
[b] Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, PR China
[c] Center for Combinatorics, Nankai University, Tianjin 300071, PR China

ARTICLE INFO

Article history:
Received 31 December 2013
Received in revised form 28 May 2014
Accepted 2 June 2014
Available online 21 June 2014

Keywords:
Face recognition
Wavelet transform
Pixel-level fusion
Feature-level fusion
Principle component analysis

ABSTRACT

The traditional wavelet-based approaches directly use the low frequency sub-band of wavelet transform to extract facial features. However, the high frequency sub-bands also contain some important information corresponding to the edge and contour of face, reflecting the details of face, especially the top-level's high frequency sub-bands. In this paper, we propose a novel technique which is a joint of pixel-level and feature-level fusion at the top-level's wavelet sub-bands for face recognition. We convert the problem of finding the best pixel-level fusion coefficients of high frequency wavelet sub-bands to two optimization problems with the help of principal component analysis and linear discriminant analysis, respectively; and propose two alternating direction methods to solve the corresponding optimization problems for finding transformation matrices of dimension reduction and optimal fusion coefficients of the high frequency wavelet sub-bands. The proposed methods make full use of four top-level's wavelet sub-bands rather than the low frequency sub-band only. Experiments are carried out on the FERET, ORL and AR face databases, which indicate that our methods are effective and robust.

1. Introduction

Face recognition, as a biological feature recognition, has been one of the most active research areas in computer vision, pattern recognition and biometrics [1]. Face recognition has several advantages over other biometric modalities such as fingerprint and iris. Besides being natural and nonintrusive, the most important advantage of face recognition is that it can be captured at a distance and in a friendly manner. Various face recognition algorithms have been devised in the literature. However, up to now, face recognition is still faced with a number of challenges such as varying illumination, facial expression and poses [2–4].

Feature extraction and classification are two key steps of a face recognition system. Feature extraction provides an effective representation of face images to decrease the computational complexity of the classifier, which can greatly enhance the performance of a face recognition system; while classification is to distinguish those features with a good classifier. Therefore, in order to improve the recognition rate of a face recognition system, it is crucial to find a good feature extractor and an effective classifier [5]. In this paper, we focus on feature extraction methods by using wavelet transform with the help of the classical principal component analysis (PCA) [6] and linear discriminant analysis (LDA) [7].

Wavelet transform, which is an increasingly popular tool in image processing and computer vision, has been investigated in many applications, such as compression, detection, recognition, image retrieval, among others, due to its great advantages with the nice features of space-frequency localization and multi-resolution. Researchers have developed face recognition algorithms by combining discrete wavelet transform (DWT) with other methods (see, for example, [8–13] and references therein). Through a two-dimensional DWT (2D-DWT), an image of face is transformed into two parts: a low frequency sub-band and three high frequency sub-bands. On one hand, the low frequency sub-band plays a dominant role in four sub-bands for approximation of the original image; and on the other hand, when the expression changes or the image is affected by occlusion, the low frequency sub-band has no obvious change but the high frequency sub-bands change obviously, so the low frequency sub-band is usually single-handed used for identification in image recognition [10]. The authors in [14] directly used the low frequency sub-band of wavelet transform to extract facial features. However, although the high frequency sub-bands do not contain

* Corresponding author at: Department of Mathematics, School of Science, Tianjin University, Tianjin 300072, PR China. Tel./fax: +86 22 27403615.

E-mail address: zt69069882@tju.edu.cn (T. Zhang).

http://dx.doi.org/10.1016/j.inffus.2014.06.001
1566-2535/

as much information as the low frequency sub-band, they also contain important information corresponding to the edge and contour of face, reflecting the details of face [15]. Moreover, we implement the classical DWT + PCA on the ORL database by using 4-level wavelet transform and every top-level's high frequency sub-band is single-handed used for identification in face recognition. The obtained recognition rates by three top-level's high frequency sub-bands are 67.5%, 61.0% and 43.0%, respectively. This implies that every top-level's high frequency sub-band contains much information of face features. Thus, it is more reasonable to use the information of more sub-bands for face recognition. A natural question is *how to use the information of these sub-bands effectively for face recognition?*

In recent years, data fusion has been developed rapidly and widely applied in many areas such as object recognition, pattern classification, image processing, and so on. Generally speaking, data fusion is performed at three different processing levels according to the stage at which the fusion takes place: pixel-level, feature-level and decision-level [16]. In pixel-level fusion, the information derived from multiple feature sets is assimilated and integrated into a final decision directly. Many fusion algorithms for the pixel-level fusion have been proposed, from the simplest weighted averaging to more advanced multiscale methods. There are two existing feature-level fusion methods. One is to group two sets of feature vectors into a union-vector, and another one is to combine two sets of feature vectors by a complex vector. Both feature fusion methods can increase the recognition rate. The advantage of feature-level fusion lies in two aspects: firstly, it can derive the most discriminatory information from original multiple feature sets; secondly, it is able to eliminate the redundant information resulting from the correlation between distinct feature sets. The decision-level fusion, delegated by multi-classifier combination, has been one of the hot research fields on pattern recognition, and has achieved successful application in face recognition.

In this paper, based on 2D-DWT, we propose a joint of pixel-level and feature-level fusion at the top-level's wavelet sub-bands technique which can make full use of four top-level's wavelet sub-bands, and we abbreviate it as TWSBF. The proposed technique is different from those image fusion techniques developed in the literature (see, for example, [17–20] and references therein). Firstly, we conduct dimension reduction on the low frequency wavelet sub-band. Then, we consider the top-level's three high frequency sub-bands, in which we apply the pixel-level fusion. In this way, we can keep more discriminatory features and avoid redundancy. We convert the problem of finding the best fusion coefficients to two optimization problems based on PCA and LDA, respectively; and propose two alternating direction methods to solve the corresponding optimization problems for finding the optimal transformation matrices of dimension reduction and optimal fusion coefficients of wavelet sub-bands. Finally, we process feature-level fusion on the low frequency sub-band after dimension reduction and fused high frequency vector. The experiments are carried out on the FERET, ORL and AR face databases. The numerical experimental results demonstrate that the proposed methods possess higher recognition rates than the classical wavelet-based methods and some popular methods at present.

The rest of this paper is organized as follows. In Section 2, after a brief introduction of 2D-DWT, we propose the model of TWSBF. The model and algorithm based on TWSBF and PCA are investigated in Section 3; and the model and algorithm based on TWSBF and LDA are discussed in Section 4. The numerical experimental results on the FERET, ORL and AR face image databases are reported in Section 5. The final conclusions are given in Section 6.

2. Model of TWSBF

Let $L^2(R)$ denote the square integrable space. The continuous wavelet transform of a one-dimensional function $f \in L^2(R)$ is defined as

$$W_f(a,b) = \int_R f(t)\overline{\psi_{a,b}(t)}dt, \tag{1}$$

where the wavelet basis functions $\psi_{a,b}(t)$ can be expressed as

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right),$$

in which $\psi(t)$ is called mother wavelet and the parameters a and b stand for the scale and position, respectively. Eq. (1) can be discretized by imposing restrictions on a and b with $a = 2^n$ and $b \in Z$.

The DWT of a one-dimensional signal is processed by transforming it into two parts with a low-pass filter and a high-pass filter. The low frequency part is split again into two parts of high and low frequencies [21]. In image processing, similar to the one-dimensional DWT, the DWT for a two-dimensional image can be constructed by applying the one-dimensional DWT at horizontal and vertical directions, respectively. The process of 2D-DWT is shown in Fig. 1, in which an image is first filtered in the horizontal direction with low-pass and high-pass filters. Then the filtered outputs are downsampled by a factor of 2. Moreover, the same process is applied in the vertical direction. Thus, an image is decomposed into 4 sub-bands denoted by LL, HL, LH, HH. Each sub-band can be thought of a smaller version of the image representing different properties. The sub-band LL is a coarser approximation to the original image; the sub-bands HL and LH record the changes of the image along horizontal and vertical directions, respectively; and the sub-band HH shows the changes of the image along diagonal direction. The sub-bands HL, LH and HH are all the high frequency components of the image. Further decomposition can be conducted on the LL sub-band. The 1-level 2D-DWT means that an original image is decomposed into a low frequency sub-band and three high frequency sub-bands. Generally, the t-level 2D-DWT $(t > 1)$ means that the low frequency sub-band obtained in the $(t-1)$-level 2D-DWT is further decomposed into a low frequency sub-band and three high frequency sub-bands. Fig. 2(a) shows the flowchart of a 2-level 2D-DWT and Fig. 2(b) shows the face image with a 1-level 2D-DWT.

Generally, sub-band LL is the low frequency component of the image which contains most information of the original image; and sub-bands HL, LH and HH stand for the high-frequency components of the image which reflect the details of images. Since the low frequency sub-band plays a dominant role in four sub-bands for approximation of the original image, if a t-level wavelet-based method is used, then the low frequency sub-band obtained by the t-th-level wavelet transform (i.e., the top-level's low frequency sub-band) is usually single-handed used for identification in image recognition in the literature. However, although the high frequency sub-bands do not contain as much information as the low frequency sub-band, they also contain much important information. Thus, it is more reasonable to use the information of all top-level's sub-bands for face recognition. Based on these analysis, we now propose a new wavelet-based technique (TWSBF) for feature extraction, which makes full use of all top-level's wavelet sub-bands. Such a technique is described as follows.

Let $I(x,y)$ be a face image, by using 1-level or multi-level wavelet transform to $I(x,y)$, we get the top-level's low frequency sub-band matrix $LL \in R^{m\times n}$ and high frequency sub-bands matrices $HL, LH, HH \in R^{m\times n}$. Without loss of generality, let $L, H, V, D \in R^l$ represent the vectorization of LL, HL, LH, HH, where $l = m \times n$. Suppose