Optik 127 (2016) 2670–2678

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Expression invariant face recognition using local binary patterns and contourlet transform

Hemprasad Y. Patil*, Ashwin G. Kothari, Kishor M. Bhurchandi

Department of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, Nagpur 440010, Maharashtra, India

ARTICLE INFO

Article history:
Received 4 September 2014
Accepted 25 November 2015

Keywords:
Face recognition
Expression invariance
Contourlet transform
Image formation and processing
Image analysis

ABSTRACT

Face recognition is one of the most widely used biometric techniques in surveillance and security. Although many state-of-the-art systems have already been deployed across the world, the main challenge for feature extraction comes from the variations in the query images captured under uncontrolled situations. Unlike the local binary patterns or steerable pyramids which construct a feature vector strictly from spatial and transform domain, respectively, our approach built a method that exploits the features from both spatial as well as contourlet transform domain. Specifically, the contourlet transform exhibits properties like directionality and anisotropy and hence, results in extraction of significant features. Furthermore, we have proposed a novel coefficient enhancement algorithm which is applicable on the contourlet subbands to make the system more robust by enhancing skin region features. In addition, we show that the feature level fusion produces a robust feature vector, which yields competitive face recognition rates on the Cohn–Kanade (CK), Yale, JAFFE, ORL, CMU-AMP and our own face database. Finally, we benchmark our approach with other contemporary approaches and found it as most robust expression invariant face recognition technique.

1. Introduction

Face recognition lies at the heart of most of the biometric systems due to its non-intrusive behavior. Some of the classical approaches for face recognition are PCA [1], LDA [2] and ICA [3]. Currently the research focus has been shifted from holistic methods towards local pixel level feature extraction techniques. It is assumed that more robust face recognition systems can be designed using the local features that are extracted with dense descriptors [4]. Few researchers have implemented the local pixel level feature extraction based systems [4,5]. Although such systems achieve higher recognition rates than holistic feature extraction based approaches, the performance of such systems hampers due to variations like expression, pose, illumination and occlusion changes in the query image.

Facial expression variations are resulting from various emotions or verbal communications. Facial expressions are helpful for human-computer interaction (HCI) devices but act as a challenge for face recognition systems. Expression invariant face recognition problem is a prominent challenge due to drastic changes in a query image arising from facial deformations. As query images have different expressions than the training images, robust feature extraction methods play a vital role to improve the recognition rates of expression invariant face recognition systems. A thorough review of 2-D face recognition literature can be found in [6]. Expression invariant face recognition approaches can be classified into two groups: (1) Spatial domain techniques (2) Transform domain techniques. The spatial domain techniques are further divided into two categories: (i) Model based techniques and (ii) Subspace based techniques. The methods based on model based techniques employ appearance based models for facial fitting. These models utilize shape as well as texture features extracted from a query face image. Riaz et al. [7] implemented an active appearance model (AAM) based framework along with optical flow based feature extraction. The composite vector constructed from the AAM parameters and optical flow results into a robust facial image representation. Lee and Kim [8] extracted a feature vector from query image by fitting AAM. Further, they employed expression state analyzer which recognizes the expression and subsequently transform the initial feature vector into neutral facial feature vector. Although these techniques achieve promising results, fitting AAM is computationally heavy. Subspace based techniques utilize the traditional statistical procedures to analyze different attributes. Li et al. [9] implemented the notion of detached texture and geometric facial information. Both texture and geometric information are projected onto distinct principal component analysis spaces to further extract

* Corresponding author. Tel.: +91 9096019335.
E-mail address: hemprasadpatil@gmail.com (H.Y. Patil).

http://dx.doi.org/10.1016/j.ijleo.2015.11.187
0030-4026/

the relevant features. Chen and Lovell [10] employed principal component analysis (PCA) and subsequently wrap the constructed subspace to enhance the class separation performance. For adaptive subspace construction, they have performed whitening and filtering operations on the scatter matrix.

Transform domain based techniques extract the eminent features by transforming images into another domain. Vankayalapati and Kyamakya [11] extracted features from combination of the wavelet transform and Radon transform. Since the Radon transform is effective for extraction of local features and wavelet transform yields the time-frequency localization; their combination results into a robust feature vector for expression invariant face recognition. Abbas et al. [12] proposed a preprocessing technique to address expression variation problem. An original facial image in transformed into wavelet domain and detailed coefficients are equated to zero. Further, upon inverse transformation, they have obtained an image representation which is robust for expression variations. Kirtac et al. [13] computed the Gabor wavelet transform and performed nearest neighbor discriminant analysis. Gabor wavelet based features yield promising recognition rates as their kernels have similarity with the human visual cortex fields. To address the issue of more robust feature extraction, multi-directional transforms like contourlet transform [14] and curvelet transform [15] can be utilized. As contourlet transform is entirely proposed in digital domain [14], it does not suffer from critical sampling. Following approaches utilize the contourlet transform for facial feature extraction. Yi et al. [16] employed the threshold based algorithm on contourlet coefficients and performed classification of facial images using the SVM classifier. Boukabou and Bouridane [17] extracted pixel level features using contourlet transform and performed dimensionality reduction using PCA. Chen et al. [18] extracted the features from low frequency coefficients using statistical parameters. Histograms of high frequency contourlet coefficients of facial images are also exploited for efficient feature extraction. Weighted similarity measure is used for matching amongst various features. Pingfeng et al. [19] utilized pixel level features from contourlet transform coefficients and performed dimensionality reduction using coupled subspace analysis. All of these techniques do not include features obtained with dense pixel-level analysis.

We propose a novel feature level fusion technique that considers features from spatial as well as contourlet transform domain. The overall flow of proposed expression invariant face recognition system is presented in Fig. 1. Main contribution of this paper is to propose a novel boosting function for contourlet coefficient enhancement and feature level fusion. We have explored our approach on databases that contain drastic expression variations like the Cohn–Kanade (CK), Yale, JAFFE, ORL, CMU-AMP database

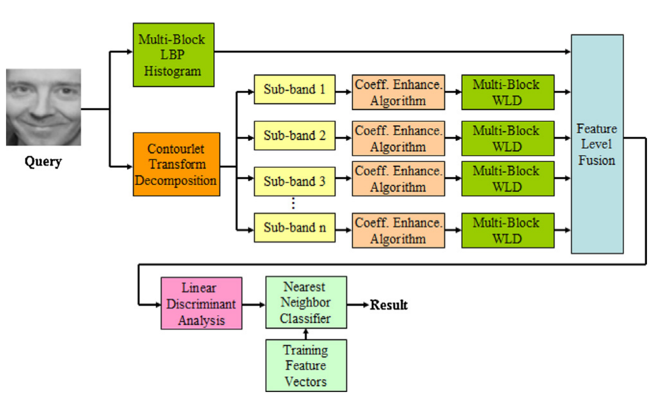

Fig. 1. Overall flow of the proposed face recognition method.

and few of our own acquired images. The proposed algorithms are equally efficient with all the databases.

The outline of the rest of the paper is as follows. In Section 2, contourlet transform, Local binary patterns and Weber local descriptors are explained. Further, a novel coefficient enhancement function is proposed in Section 3. In Section 4, a feature extraction and feature level fusion scheme is presented along with feature matching using nearest neighbor technique. Experimentation on various face databases and results are discussed in Section 5.

2. Materials and methods

2.1. Contourlet transform

Most of the classical frequency domain transforms such as Fourier transform, Discrete Cosine Transform and Wavelet transform are very efficient in capturing the details as far as they are exploited in one-dimensional context. In order to extract the details of an image, 2-D extensions are constructed from the 1-D separable basis functions. Such transforms are good at capturing the details when the image is a collection of 1-D piece-wise smooth signals [14]. However, practical applications like face recognition, human computer interaction demand the processing of images that involve smooth curves. For example, facial images have smooth curves that represent eyes, eyebrows, lips, nostrils etc. These smooth curves contain discriminant information which is essential for face recognition. Wavelet like transforms cannot effectively capture the features from contour like edges [14]. To address the problem of efficient representation of contour-like smooth edges, Do and Vetterli [14] proposed the contourlet transform directly in the discrete domain. Contourlet transform can successively approximate the image to have multiresolution representation. It can critically sample the original image with a small redundancy. It also represents the localization in spatial as well as frequency domain. In addition, its basis elements can be aligned to many directions which is useful to extract the directional features. Notably, its basis elements can be extended to form the shape of smooth curves. Although 2-D Gabor wavelet transform [20] and contourlet transform both can provide multidirectional and multiresolution expansions, the latter maintains critical sampling with small redundancy. The contourlet transform decomposition consists of the following stages: First, the original image is subjected to modified Laplacian pyramid (LP) transform [21] which acts as a wavelet like transform and hence utilized for edge detection. Further, on the resultant LP decomposition, modified iterated directional filter bank [22] is employed to extract smooth contour segments [14]. The modified Laplacian pyramid transform and modified iterated directional filter banks are illustrated in the following sub-sections.

2.1.1. Modified Laplacian pyramid transform

Burt and Adelson [23] proposed the Laplacian pyramid transform for image coding. At each level of the LP decomposition, one low pass downsampled image (say *LPD*) is constructed by passing the original image through low pass analysis filter $H(w)$ and further downsampling it with factor N. In addition, image *LPD* is upsampled with factor N and filtered with synthesis filter $G(w)$. This results in the image *HPD*. The difference between original image and *HPD* is encoded as bandpass image (*BP*). This decomposition procedure is illustrated in Fig. 2. The *LPD* image is further subjected to similar decomposition procedure based on number of levels of decomposition.

Do and Vetterli [21] proposed the modified Laplacian pyramid decomposition by altering the analysis and synthesis filters to orthogonal filters. A series of mappings $\{\Theta_k\}_{k\in\Gamma}$ defined in the bounds of Hilbert space H are denoted by a 'frame' if the following