Neurocomputing 177 (2016) 49–61

Contents lists available at ScienceDirect

Neurocomputing

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

Non-stationary feature fusion of face and palmprint multimodal biometrics

Muhammad Imran Ahmad [a,*], Wai Lok Woo [b], Satnam Dlay [b]

[a] School of Computer and Communication Engineering, Universiti Malaysia Perlis, Perlis, Malaysia
[b] School of Electrical and Electronic Engineering, Newcastle University, Newcastle upon Tyne, United Kingdom

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form
19 September 2015
Accepted 3 November 2015
Comminicated by Jungong Han
Available online 3 December 2015

Keywords:
Multimodal biometrics
Feature fusion
Information fusion
Local features
Face recognition
Palmprint recognition

ABSTRACT

Multimodal biometrics provides high performance in biometric recognition systems with respect to unibiometric systems as they offer a more universal approach, added security and better recognition accuracy. Moreover, data acquisition at the feature level brings out rich information from the traits, thus fusion of modalities at this level is desirable. In this paper we propose a novel fusion technique called non-stationary feature fusion where a new structure of interleaved matrix is constructed using local features extracted from two modalities i.e. face and palmprint images. A block based Discrete Cosine Transform (DCT) algorithm is used to construct a fused feature vector by extracting independent feature vectors from each spatial image. This fused feature vector contains nonlinear information that is used to train a Gaussian Mixture Models (GMM) based statistical model. The model provides accurate estimation of the class conditional probability density function of the fused feature vector. The proposed method produces recognition rates as high as 99.7% and 97% when tested on benchmark databases-ORL-PolyU and FERET-PolyU respectively. These rates are achieved using 23% low frequency DCT coefficients. The new technique is shown to outperform existing feature level fusion methods including methods based on matching and decision level fusion.

1. Introduction

BIOMETRIC recognition is well established in important areas of law enforcement and control of physical, health and logical access. Biometric identification manipulates unique behavioral or physiological traits to determine the identity of an individual [1]. Fingerprint [2], iris [3], face [4] and palmprint [5,44] are among the most widely used traits as these biometric algorithms are exclusive. A single trait biometrics is known as single modal system, whereas those that integrate two or more traits are classified as a multimodal system. In most cases, the practice of using a single modal biometrics is inadequate for individual identification due to several problems such as noise, intra-class variations, restricted degree of freedom, non-universality, spoof attacks and unacceptable error rates [6].

Therefore, a robust identification system requires several modalities to address the limitations in single modal biometrics. The multimodal system utilizes a fusion approach to combine the information obtained from various individual modalities hence it has the capability to perform better than single modal biometrics.

The information obtainable by multiple traits can be consolidated at three levels of fusion [8]: feature level [9–11]; matching score level [14]; and decision level [16]. Fusion at the feature level is performed by concatenating two or more features into a new single feature which is then used in the matching and decision module of the system. In matching score level fusion, the matching scores output by multiple matchers are integrated by using several integration methods such as sum-score, max-score and min-score [18,19]. Since the matching score output from different modalities comes in different ranges, a score normalization method (e.g. min–max, median, sigmoid, Tanh and *z*-score) is used to transform the scores into a common domain. In decision level fusion, the final decision made by individual systems is consolidated by employing techniques such as a majority vote, AND rule and OR rule [20,21]. Fusions at the matching score level and decision level are commonly used in practice since they are more easily implemented. However, fusion at the feature level is believed to produce better recognition output despite being used less often than the other two methods. The strength of feature fusion technique lies its ability to, derive highly discriminative information from original multiple feature sets; and to eliminate redundant information that results from the correlation between

* Corresponding author.
E-mail addresses: m.imran@unimap.edu.my (M.I. Ahmad), lok.woo@ncl.ac.uk (W.L. Woo), satnam.dlay@ncl.ac.uk (S. Dlay).

http://dx.doi.org/10.1016/j.neucom.2015.11.003
0925-2312/

distinct feature sets, thus gaining the most effective feature set with low dimensionality for the final decision.

Most of the present researches referring to the fusion of biometrics at the feature level use concatenation methods. Concatenation can be performed after feature extraction such as in the concatenation of magnitude information of Gabor images [9,11]. The Gabor images representation captures salient visual properties such as spatial localization and orientation thus superior performances are reported. The major drawbacks of this method are in terms of high dimensionality and computational complexity due to the use of Gabor filter bank with different scales and orientations. Fused Gabor images that contain non-linear information are further processed to extract important information as well as to reduce their dimension using linear [11] and nonlinear dimensional reduction methods [9]. The nonlinear reduction method [24,44] showed better performance than linear method [22,23] as biometric images are normally embedded with non-linear information, The only drawback is that it is computationally more demanding.

Instead of concatenating information after feature extraction, several methods [10,11,25] concatenate features after a dimensional feature reduction using several approaches such as principal component analysis (PCA), linear discriminant analysis (LDA) and independent component analysis (ICA). Yao [11] and Ahmad [10] used PCA and LDA to extract information from two modalities and merged the feature vectors to form a long 1-D vector. However, these linear feature reduction methods did not fully utilize the inherent non-linear information that exists in the modalities thus producing a merged feature vector with less discriminant power [24]. Concatenation can also be undertaken on raw data (i.e. without pre-processing) by using a feature extraction method [25], where all the raw data was normalized to the same size and sorted in a long vector prior to the application of a subspace selection method to construct the transformation matrix. A new subspace selection method was utilized in this method, based on linear discriminant algorithms. This method offered better computational speed, but limited to raw data that is not highly non-linear. To overcome the nonlinearity problem in raw data, a kernel trick was adopted to map raw data to high dimensional feature space.

To date, most of the feature level fusion methods use concatenation process, which serially combine features from two or more modalities to form a long vector. In order to extract important information and reduce the dimensionality, linear and nonlinear feature extraction methods are used prior to and after the fusion process. By using a concatenation process that does not take into account the distribution of data in both modalities, some data may become redundant, and thus it will not fully utilize the fusion itself. In reality, some of the modalities are nonlinearly distributed (e.g. face image with different pose and expression, palm-print image with different agieng), thus utilizing linear subspace reduction techniques cannot fully exploit the information contained in these modalities.

In this paper, we propose a matrix interleaved feature fusion that is able to extract and capture non-stationary features from both modalities especially those having different statistical properties. Non-stationary information in biometric data occurs due to variation in pose and ageing. The proposed method can be summarized as follows:

- Introduction of a new structure of feature fusion using matrix interleaved concatenation method. The new structure will change the distribution of the structure matrix by interleaving the extracted face and palm features to produce features with a different statistical distribution. Compared to conventional methods which apply concatenation methods to merge feature vectors in order to form a single vector for classification, the present method is designed to estimate model parameters of the concatenation features and capture their statistical properties. Further, fusion using our method has the benefit of having large data points compared to the conventional concatenation methods that use less data points. Low dimensional feature vector is concatenated in a matrix interleaved producing double data point and double dimension. An accurate statistical model can be estimated by using large number of fused feature vector [41].
- The proposed method makes use of low frequency coefficients extracted using the DCT transform to represent local features prior to combining them in the matrix interleaved framework. Local features of each face and palmprint region are extracted from 8×8 sub windows with several pixels overlapping among each window.
- Model parameters of a new feature vector distribution are estimated using parameterized statistical model based on GMM. The advantage of GMM is that different types of fused data distributions can be represented as a convex combination of several normal distributions with respective means and variances. Various mixture components (i.e. weight, mean and variance) of the data distribution are calculated using maximum likelihood (ML) estimation [26]. Expectation Maximization (EM) algorithm is used to find the maximum likelihood estimation of the parameters of an underlying distribution from a given non-stationary data.

The feasibility of the proposed method has been successfully tested on two multimodal datasets created from face and palm print biometrics (1) FERET-PolyU database (2) ORL-PolyU database. The effectiveness of the proposed method is demonstrated in terms of comparative performance against the other multimodal biometrics using feature level fusion. The rest of the paper is organized as follows: Section 2 presents our proposed fusion method; Section 3 details the experimental results and discussion; and Section 4 sets out the overall conclusions.

2. Framework of the proposed method

The proposed fusion method integrates information of face and palmprint images at feature level. A new structure of matrix interleaved and statistical model is used to capture the underlying statistical information in the fused feature vector as shown in Fig. 1. The advantages of the proposed method are as follows:

1) The main features in the modalities are extracted using DCT transform and represented in different frequency bands. Local features extracted in sub-block windows consist of information that varies from low to high frequency coefficients [42]. Integration of such information can increase the discrimination power in the fused feature space.
2) The use of GMM as a statistical tool enables modeling of non-linear distribution of fused features. GMM allows huge and high dimensionality fused features to be represented in a low dimensional feature vector given by a set of GMM components (i.e. weight, mean and covariance). By increasing the number of GMM components, most of the non-stationary information exists in the modalities can be accurately learnt.
3) Matrix interleaved fusion using low frequency coefficients can eliminate the problem associated with high dimensionality features which is a common issue in concatenation methods. Fusing performed in this work uses local features extracted in small sub windows thus production of high dimensional features can be avoided. Nonlinear boundary problem is solved by using GMM to capture any shape of feature distribution in the fused feature space.