Int. J. Electron. Commun. (AEÜ) 67 (2013) 733–742

Contents lists available at SciVerse ScienceDirect

International Journal of Electronics and Communications (AEÜ)

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

Adaptive score level fusion of fingerprint and voice combining wavelets and separability measures

Anzar S.M.*, Sathidevi P.S.

Department of Electronics and Communication Engineering, National Institute of Technology, Calicut 673601, India

ARTICLE INFO

Article history:
Received 7 September 2012
Accepted 3 March 2013

Keywords:
Biometrics
Integration weight
Noise robustness
Score denoising
Separability measures

ABSTRACT

This paper presents an adaptive combinational approach for the score level fusion of fingerprint and voice biometrics, whose performance under adverse noise conditions are investigated systematically. An efficient preprocessing on the raw vector of scores using normalization and wavelet denoising is proposed, to improve the performance of the multibiometric system. The class as well as the score separability measures, under various noise conditions are estimated and combined algebraically, to determine the best integration weights, for the complementary modalities employed. The *z*-score normalized impostor scores are modelled as white Gaussian noise so that the wavelet denoising techniques can be readily applied. The inter/intra class separability measures from the feature space and the d-prime separability measures from the match score space are estimated in the training/validation phase. The performance of the proposed method is compared with the baseline techniques on score level fusion. Experimental evaluations show that the proposed method improves the recognition accuracy and reduces the false acceptance rate (FAR) and false rejection rate (FRR) over the baseline systems, under various signal-to-noise ratio (SNR) conditions. The proposed biometric solutions will be extremely useful in applications where there are less number of available training samples.

1. Introduction

Noise robustness of a multimodal system is still an unsolved issue. Multibiometric recognition under noisy environment relies on the data coming from unimodal sources. When a biometric measure obtained from one modality is corrupted by noise, the evidence presented by a comparatively reliable trait may be used for more accurate determination of identity [1,2]. Biometric indicators such as fingerprints, iris, and retina, have high degree of permanence [3] while the traits such as signature, face and voice have high degree of variance. Although the recognition accuracy of the voice biometric is high in clean conditions, its performance tends to be significantly degraded under the presence of background noise. The drawback of the current fusion techniques is the inability to cope up with varying environmental conditions such as sensor noise and ageing factors [3]. Moreover, the estimation of the best integration weight is important as it determines the amount of contribution of each modality towards the final decision, otherwise, the system may perform attenuating fusion. The focus here is to consider the intelligent information fusion, combining the class separability and score separability measures using score normalizations and wavelet based score preprocessing techniques. Haar wavelet functions are used to denoise the raw matching score vectors.

The inter/intra class separability measures derived from the feature space and the d-prime separability measures from the match score space are estimated separately for each noise condition in the training/validation phase. The performance of the proposed scheme has been compared with that of equal weight bimodal biometric systems, grid search (GS) and genetic algorithm (GA) based optimal integration schemes [4] and the reliability and separability based optimal integration schemes [5]. The experimental results discussed in Section 6 show that, the proposed method results in robust recognition accuracy even under low SNRs. The rest of the paper is structured as follows. The following section gives a brief survey on the related work available in literature. Section 3 gives a brief outline of the individual classifiers and matching strategies used. Section 4 describes the proposed score normalization and wavelet based matching score preprocessing techniques. Section 5 discusses the estimation of the best integration weights using separability measures. Results of the experiments are detailed in Section 6. The paper concludes with a brief summary in Section 7.

2. Related work

Bengio et al. proposed a method to integrate the confidence measures as weights for multimodal fusion [6]. Toh proposed a generalized reduced multivariate polynomial model for combining

* Corresponding author. Tel.: +91 9447244119; fax: +91 4952287250.
E-mail addresses: p090004ec@nitc.ac.in (Anzar S. M.), sathi@nitc.ac.in (Sathidevi P. S.).

1434-8411/$ – see front matter © 2013 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.aeue.2013.03.004

fingerprint and speaker verification decisions [7]. Lewis et al. shed some light on audio-visual speech recognition systems using dispersion measures as the integration weights [8]. Toh et al. combined fingerprint and speaker verification decisions in the match score level using functional link network [9]. Poh et al. proposed a margin derived confidence measure while fusing two system opinions [10]. Kryszczuk et al. proposed a method of performing multimodal fusion using face and speech data combining signal quality measures and reliability estimates [11]. Morizet et al. proposed an adaptive combinational approach to score level fusion for face and iris biometrics combining wavelets and statistical moments [12]. Alsaade et al. showed that score normalization and quality-based fusion improves the accuracy of multimodal biometrics [13]. Optimal integration weight estimation using least squares technique was reported in [7]. Integration weight optimization for fingerprint and voice biometrics based on grid search and genetic algorithm was reported in [14]. Reliability based optimal integration weight estimation was reported in [4,5]. Multibiometric authentication using discrete wavelet transform (DWT) and score level fusion was presented by [15]. Here, our objective is to develop a bimodal system, with fingerprint and voice biometrics, that is more robust to environment and sensor noise. We have determined the *best integration weight* β combining the separability measures derived from the feature space and the matching score space using score normalization and wavelet based score denoising techniques. To the best of our knowledge, the proposed adaptive score level fusion of fingerprint and voice combining wavelets and separability measures have not been attempted until now.

3. Individual classifiers

3.1. Fingerprint classifier

We have considered the minutiae-based fingerprint matching algorithm using ridge counting [16]. Each minutiae is represented as a triplet $m = \{x, y, \theta\}$ that indicates the x, y minutiae location coordinates and the minutiae angle θ. A minutiae m_i in T and a minutiae $m_j^{'}$ in I are considered matching, if the spatial distance (sd) between them is lesser than a given tolerance r_0 and the direction difference (dd) between them is lesser than an angular tolerance θ_0 [17].

$$sd(m_j^{'}, m_i) = \sqrt{(x_j^{'} - x_i)^2 + (y_j^{'} - y_i)^2} \leq r_0 \tag{1}$$

$$dd(m_j^{'}, m_i) = min(|\theta_j^{'} - \theta_i|, 360^0 - |\theta_j^{'} - \theta_i|) \leq \theta_0 \tag{2}$$

Elastic matching algorithm is used to perform matching between the two fingerprints. Match score formula for the reference and the test print is given by [17],

$$\text{Matching score} = \frac{100 N_{\text{pair}}}{max\{M, N\}} \tag{3}$$

where N_{pair} is the number of matched minutiae, M is the number of minutiae in the template set, and N is the number of minutiae in the test set. Maximum similarity criterion is used for fingerprint pattern classification.

3.2. Voice classifier

Short-time spectral analysis is used to characterize the quasi-stationary speech samples. To represent the voice samples in a parametric way, we have considered the cepstral representation as it has been proved to be efficient and compact [18]. The number of mel cepstrum coefficients, is chosen as 16 (here). Gaussian mixture model (GMM) is considered here for representing the acoustic feature vectors. The complete GMM is parameterized by the mean vectors, covariance and the mixture weights. These parameters are collectively represented by [19],

$$\lambda = \{a_i, \mu_i, \Sigma_i\}, \quad i = 1, \ldots, M \tag{4}$$

In the training stage itself, each enrolled speakers in **g**, (where $\mathbf{g} = \{\hat{g_1}, \hat{g_2}, \ldots, \hat{g_G}\}$) is represented by a unique GMM (λ). In the testing stage, the features from the unknown speaker's utterances are compared with statistical models of the voices of speakers known to the system. The Bayes rule allocates the test samples to the class $\hat{g_k}$, having the highest posterior probability, that is [19],

$$\hat{g_k} = arg \max_{1 \leq k \leq G} p(X|\lambda_k) \tag{5}$$

where $p(X|\lambda_k)$ is the *a posteriori* probability for a given observation sequence.

4. Matching score preprocessing techniques

Environmental noise and the quality in acquisition affect the matching score distribution. If the matcher module is not able to factor out these peculiarities in the acquisition process, the match scores are affected by these environmental variations. This may increase the inter-class similarity scores and decrease the intra-class scores. Here, we have considered the score preprocessing techniques such as normalization and wavelet based score denoising techniques. The matching score normalization techniques effectively normalizes any unwanted peculiarities involved in the raw similarity computations [20] while wavelet denoising of matching scores keeps the genuine scores as high as possible and keeps the impostor scores as low as possible [12] thereby enhancing the robustness as well as the efficiency of the recognition system. Score preprocessing techniques are applied in both the training and the testing phase.

4.1. Score normalization

The central idea behind score normalization is to reduce the data variations that are reflected in the matching scores. This transformation, essentially re-allocates the location and scale parameters of the score distributions [20]. Location parameter simply shifts the distribution curve left or right on the horizontal axis. The effect of the scale parameter is to stretch or compress the distribution curve. For the normal distribution, the optimal location and scale parameters correspond to the mean and standard deviation, respectively. For an arbitrary distribution, mean and standard deviation are reasonable estimates of location and scale parameters, respectively, but are not optimal. Various score normalization techniques are proposed in the literature. For a good normalization scheme, the estimates of the location and scale parameters of the matching score distribution must be robust and efficient [20]. We have considered z-score and tanh normalization techniques.

Score normalization is defined as a function that maps s_i to $s_i^{'}$, where $s = \{s_i\}, i = 1, 2, \ldots, n$, is the set of matching score vectors and $s_i^{'}$ is the normalized scores [21]. The parameters of the normalization techniques are obtained from the genuine and impostor matching score distributions that are generated on the training data. During training, the set s_i is divided into two subsets, s_i^G and s_i^I, which denote the raw genuine and imposter matching scores, respectively. The *z*-score (ZS) normalization method transforms the scores to a distribution with zero mean and unit standard deviation. The normalized scores are given by

$$s_i^{'} = \frac{s_i - mean(s_i^G, s_i^I)}{std(s_i^G, s_i^I)} \tag{6}$$