Information Fusion 32 (2016) 13–25

Contents lists available at ScienceDirect

Information Fusion

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

Towards a multi-source fusion approach for eye movement-driven recognition

Ioannis Rigas*, Evgeniy Abdulin, Oleg Komogortsev

Department of Computer Science, Texas State University, San Marcos, TX 78666, USA

ARTICLE INFO

Article history:
Available online 20 August 2015

Keywords:
Eye movement biometrics
Multi-stimulus fusion
Multi-algorithmic fusion

ABSTRACT

This paper presents a research for the use of multi-source information fusion in the field of eye movement biometrics. In the current state-of-the-art, there are different techniques developed to extract the physical and the behavioral biometric characteristics of the eye movements. In this work, we explore the effects from the multi-source fusion of the heterogeneous information extracted by different biometric algorithms under the presence of diverse visual stimuli. We propose a two-stage fusion approach with the employment of *stimulus-specific* and *algorithm-specific* weights for fusing the information from different matchers based on their identification efficacy. The experimental evaluation performed on a large database of 320 subjects reveals a considerable improvement in biometric recognition accuracy, with minimal equal error rate (EER) of 5.8%, and best case Rank-1 identification rate (Rank-1 IR) of 88.6%. It should be also emphasized that although the concept of *multi-stimulus fusion* is currently evaluated specifically for the eye movement biometrics, it can be adopted by other biometric modalities too, in cases when an exogenous stimulus affects the extraction of the biometric features.

1. Introduction

The human body provides an invaluable source of distinctive information suitable to be used for the task of biometric recognition [1]. The most well-studied and widely-adopted biometric modalities are the fingerprints, the iris, and the face. Some other explored biometric traits include the palm, the hand geometry, the ears, the nose, and the lips. The analysis of the blood-vessels morphology appears as the main source of biometric features in methods like the vein matching, and the retinal scan. There are also some biometric traits that enfold behavioral characteristics, i.e. traits that are partially connected with the brain activity. Examples of this category involve the speech analysis and voice recognition, the hand-written signature, the keystroke dynamics, the gait analysis, and the eye movement-driven biometrics. Considering the abundance of the existing biometric modalities and the heterogeneity of the associated features, it may come as no surprise that there is a strong trend in the biometric research towards the investigation and adoption of information fusion techniques.

1.1. Information fusion in biometrics

Information fusion can provide numerous benefits in the domain of biometric recognition. The most obvious among them is the expected performance gain in terms of biometric accuracy due to the combination of evidence gathered from multiple cues [2]. Also, the fusion techniques can be employed for the selection and the promotion of the most informative features among a large set of such features [3]. In addition, the combination of different sources of biometric information can open the path for the creation of biometric systems with enhanced robustness against security flaws and spoofing attacks [4].

The fusion of biometric information can be implemented in multiple ways. A common approach is to combine the information coming from different modalities (e.g. fingerprints, face, iris etc.). An early work demonstrating such a multi-modal fusion scheme for fingerprint and face cues was presented in the work of Hong and Jain [5]. Also, one of the first important studies evaluating the information fusion of fingerprint, face, and hand geometry cues was presented by Ross and Jain in [6]. The study presented by Yang et al. [7], investigated the fusion of characteristics that can be extracted exclusively from the hand region, such as the fingerprints, the hand geometry, and the palm-prints. Analogously, several approaches focused on the fusion of information coming from the face and the head area, given the abundance of distinct characteristics of these specific body regions. In the work of Wang et al. [8], face and iris features were fused

* Corresponding author. Tel.: +1 512 245 0349.
E-mail addresses: rigas@txstate.edu (I. Rigas), e_a146@txstate.edu (E. Abdulin), ok11@txstate.edu (O. Komogortsev).

http://dx.doi.org/10.1016/j.inffus.2015.08.003
1566-2535/

in order to combine the virtues of both modalities. The study of Chang et al. [9] involved an appearance-based fusion scheme employing images of the face and the ear. Another category of multi-modal fusion techniques proposed the combination of physical and behavioral biometric cues. Voice and face were among the first combined features [2,10], whereas other scenarios involved the combination of face and keystroke dynamics [11], and face and gait features [12].

A different type of information fusion in biometrics involves the combination of the data coming from a single biometric modality by applying multi-algorithmic fusion techniques. In the field of fingerprint biometrics there are several examples of information fusion implemented using multiple algorithms in different stages of the recognition process [13–15]. The work presented by Vatsa et al. [16] employed the iris as the single modality for implementing multi-algorithmic information fusion. Different techniques for performing multi-algorithmic fusion were also evaluated for the face biometrics [17,18], in an attempt to use the variability of the features of this specific modality. In the work presented by Han and Bhanu [19], a multi-algorithmic scheme was used for the behavioral trait of gait via the analysis of the influence of the external conditions on the gait patterns.

Several multi-instance fusion techniques were developed in an effort to improve the accuracy of the single-modality biometric systems in practical scenarios. The FBI's IAFIS system [20] can capture the fingerprints of all ten fingers and combines the information for producing more accurate results, a technique proven to be particularly robust when operating on large databases. The work presented by Prabhakar and Jain [21] suggested the fusion of the impressions of multiple fingers by employing multiple (four) algorithms, thus creating a scheme for performing both multi-instance and multi-algorithmic fusion. Also, the work presented by Jang et al. [22] proposed a multi-unit fusion approach for the iris biometrics, using the images coming both from the left and from the right eye in order to address the quality issues often occurring when capturing a single instance of the iris.

Irrespectively of the use of a single or multiple modalities, the fusion methods can be also categorized with respect to their involvement in the typical processing levels (modules) followed in the biometric recognition routine [6], i.e. the sensor level, the feature level, the comparison (or matching) score level, and the decision level. Information fusion in the sensor level can be performed by using the data captured by different types of sensors, e.g. optical and capacitance sensors [23]. Fusion in the feature level can be implemented via the direct incorporation of the extracted features into a compact feature representation [9,24]. However, in several occasions, the nature of the feature vectors prohibits such an operation. The combination of information in the comparison score level is by far the most common strategy for implementing fusion in biometrics [2,5,6,25,26]. In this case, the universal accessibility of the comparison scores and the minimal influence of the features' heterogeneity act catalytically for the creation of efficient information fusion schemes. Finally, information fusion can be also performed in the classification stage either by using the identification ranking information [27], or by using the decisions regarding the identity or the validity of a verification claim [21,28].

1.2. Motivation and contribution

Eye movements are an emerging biometric modality [29], however, the reported performance still lacks the accuracy of the widely adopted modalities, such as the fingerprints and the iris. The existing performance gap can be attributed to the complicated mechanisms involved in the generation of the eye movements, which combine the physical characteristics of the internal eye structure [30], and the behavioral cues related to the brain activity and visual attention [31]. This work presents a multi-source fusion scheme for the combination of eye movement characteristics extracted by different algorithms (multi-algorithmic fusion) under the influence of different visual stimuli (multi-stimulus fusion). Multi-stimulus fusion is a novel concept inspired by the practically proven influence of different visual stimuli on different eye movement-driven biometric algorithms [32–34]. The theoretical background for performing the multi-stimulus fusion is also supported by several psycho-visual studies, which demonstrate the interrelationships between the visual stimulus and the generated eye movements [35–37].

The contribution of the current research in the field of eye movement biometrics can be summarized as follows:

(1) We introduce the concept of multi-stimulus fusion, i.e. fusion of different instances of the same modality (eye movements) under the influence of different visual stimuli.
(2) We propose a hierarchical weighted fusion scheme for the efficient combination of the comparison (matching) scores generated by the different eye movement algorithms (multi-algorithmic fusion) under the influence of diverse visual stimuli. Also, we suggest a weight-training method for the calculation of the fusion weights, which is based on the identification performance of different matchers.
(3) We present a comprehensive investigation of the combined effects from the multi-source fusion (multi-stimulus and multi-algorithmic) in the performance of the eye movement-driven biometrics. We provide an extensive analysis regarding the parameters of our model, and demonstrate the achieved performance improvement by using a large database of 320 subjects.

2. Research on eye movement biometrics

The first study on biometric recognition via the eye movements was presented by Kasprowski and Ober [38] a decade ago. It was based on the spectrum analysis of the eye movement signals, and used a randomly 'jumping' point of light as the visual stimulus. The reported false acceptance rate (FAR) was 1.36%, and the false rejection rate (FRR) was 12.59%. In the work of Bednarik et al. [39], the fast Fourier transform (FFT) was used along with the principal component analysis (PCA) for the analysis of the eye movements during the observation of a simple 'cross' stimulus. The achieved Rank-1 IR reached the value of 56%, and the simple form of fusion that was attempted failed to improve the results any further. The work of Kinnunen et al. [40] was inspired from the field of voice recognition, and analyzed the recorded eye movement signals during the observation of complex stimuli (text and video). The reported minimal EER was about 30%. In the work of Komogortsev et al. [32], a model of the internal non-visible structure and functionality of the eye was employed in order to implement the oculomotor plant characteristic (OPC) biometrics. In this case, the visual stimulus was a point of light making horizontal and vertical 'jumps', and the reported half total error rate (HTER) was 19%. The complex eye movement behavior (CEM-B) biometrics was introduced by Holland and Komogortsev in [33]. The used visual stimulus consisted of text excerpts, and the fusion of the comparison scores from the individual features led to an EER of 16.5%. An attempt to fuse the information of the OPC and the CEM characteristics was presented by Komogortsev et al. [41], showing a possible performance improvement of 30% over the single methods. In the work of Rigas et al. [42], a graph-based approach was used for comparing the spatial distributions of the eye fixations during the observation of stimulus consisting of human face images. The reported minimal EER was 30%. Face images were also used in the graph-based work of Cantoni et al. [43], where a minimal EER of 25% was reported. In the study of Yoon et al. [44], images of cognition-related dot-patterns were employed as the stimuli in a scheme that used hidden Markov models (HMM) to analyze gaze velocity features. The reported Rank-1 IR was in the range of 53–76%. The recent work of Rigas and Komogortsev