Neurocomputing 149 (2015) 526–535

Contents lists available at ScienceDirect

Neurocomputing

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

Combining biometric matchers by means of machine learning and statistical approaches

Loris Nanni [a,*], Alessandra Lumini [b], Matteo Ferrara [b], Raffaele Cappelli [b]

[a] *Department of Information Engineering, University of Padua, Italy*
[b] *Department of Computer Science and Engineering (DISI), University of Bologna, Italy*

ARTICLE INFO

Article history:
Received 8 August 2013
Received in revised form
14 July 2014
Accepted 4 August 2014
Available online 19 August 2014

Keywords:
Fingerprints
Machine learning
Fusion of matchers
FVC2006

ABSTRACT

In this paper we study both machine learning and statistical approaches for combining fingerprint matchers of the FVC2006 competition. We investigate not only which is the best fusion approach, but also the correlation among the state-of-the-art matchers for fingerprint verification and scanner interoperability of the fusion techniques. Several tests are performed on all the four FVC2006 datasets, using a leave-one-out dataset testing protocol, i.e., the training phase is conducted on the datasets not used in the testing phase, so it is possible to study the pros and cons of machine learning and statistical approaches when different scanners are used in the training and testing phases.

This work confirms that the fusion of different state-of-the-art fingerprint matchers can lead to a significant performance gain with respect to a single matcher.

1. Introduction

Authentication in several day-to-day activities, access control in restricted areas and other several applications require automated systems to establish the identity of a person. Biometric systems are probably the most effective technology to identify people and many commercial biometric systems have gained a large diffusion. However, the performance of these systems often is not adequate for the demand of robustness and high accuracy for some high-security applications. Although biometric verification systems have proven to be reliable in ideal environments, they can be very sensitive to real environmental conditions.

How to improve the performance of an existing biometric system in real environments is an interesting and widely studied problem: many researchers proposed biometric fusion, i.e., the use of multiple types of biometric data, or methods of processing, to improve the performance of biometric systems by consolidating the evidence presented by multiple biometric sources [2].

The general claim is that different biometrics show evidence of complementary and non-correlated characteristics which can be exploited at different levels of fusion: (i) sensor level, (ii) feature level, (iii) matching score level and (iv) decision level. The most widely adopted scheme is the so-called "score-level", based on the combination of systems' scores.

Moreover, biometric fusion can be divided into two categories according to the source of information [12]: (i) monomodality, if a single biometric trait is acquired and processed using different approaches and (ii) multimodality, if several biometric traits are used for person authentication (as [43] where the fusion between an iris matcher and a palmprint matcher is proposed).

It is well known in the literature that the main drawback of the monomodal biometric systems is the intraclass variations. For solving this problem the multi-modal system could be used. For obtaining a ~ 0 equal error rate it is possible to combine high performance biometric characteristics as iris and fingerprint [44]. The multi-modality fusion permits to obtain a very low equal error rate also if the performance of the matchers that built the ensemble is not excellent [43,45]. The drawback of these approaches is that different biometrics should be extracted using differ sensors; this increases the cost and not always it is possible to insert different sensors in the same device (e.g. smartphone). Moreover, multimodality has the advantage that it is harder to circumvent or forge, since it is more difficult to obtain and replicate multiple traits as compared to a single one; however, multimodality also means additional installation and operational costs (e.g., different acquisition devices, multiple algorithms) that could be unacceptable for some customers.

In this paper we propose a framework for monomodal biometric fusion based on a single acquisition device and multiple matching units. Anyway the fusion approaches proposed and tested in this work can be also used for multimodal fusion. The aim of this work is to show that even if a single matcher can be weak, or degrade its performance in presence of hard environmental conditions, different

* Corresponding author.
E-mail address: loris.nanni@unipd.it (L. Nanni).

http://dx.doi.org/10.1016/j.neucom.2014.08.021
0925-2312/

matchers can provide complementary information. This can allow increasing the robustness of the resulting combined system.

The selected biometric trait is the fingerprint, which is probably the most popular and studied trait in biometric community. Many researchers have already studied fingerprint biometric fusion, including fusion at a feature level [25], score level [26], rank level [38], multiple fingers [27] (also with the aim of fingerprint indexing [36]) and multiple impressions of the same finger [28]. Recently in [37] the authors proposed a novel modular framework for biometric fusion, including solutions for the data normalization problem. All the reported results agree in affirming the advantage of fusion fingerprint-based systems with respect to systems based on single fingerprint, single feature extraction or single matcher.

The proposed framework is designed not only to provide improved performance over the single fingerprint matcher, but also to explore the correlation among many state-of-the-art matchers for fingerprint verification and to study different learning approaches for combining fingerprint matchers.

In our opinion, the apparent strong correlation among matchers based on the same features (i.e., minutiae) could discourage the fusion of such algorithms. However, it should be noted that our experiments show that most of state-of-the-art approaches proposed in the literature and existing in commerce are not so correlated as one can think, as the enhancement approaches, minutiae extraction methods and matching algorithms can produce significantly different results. Therefore, the fusion of different matchers is worth theoretical and experimental investigation.

In this work we consider several state-of-the-art matchers for fingerprint verification which participated in FVC2006 [30,32] and we exploit fusion techniques based on machine learning.

Even if FVC2006 is quite dated, it is the last international contest in fingerprint recognition including several systems that can still be considered at the-state-of-the art since their results are comparable with those published on the FVC-onGoing website [41,42]. Therefore the performance of their fusion is of great interest for the biometric community.

The fusion approaches, according to the taxonomy proposed in [12], can be classified into three categories:

- *Transformation-based score fusion:* before the fusion, matching scores are first normalized to a common domain (e.g., in the interval [0, 1]): the scores can be equally weighed or the combination weights can be data-dependent in order to advantage the most performing matcher. This trained normalization/weighting is one of the main drawbacks of this class of approaches since an extensive empirical evaluation is needed [3-5,29].
- *Classifier-based score fusion:* scores obtained from multiple matchers are managed as a feature vector and used to train a second level classifier that discriminates between genuine and impostor users [2,6,7]. The main drawback of this class is that it requires a training set and often a quite large number of features [12].
- *Density-based score fusion:* this class of approaches is the most complex. It is based on the likelihood ratio test and it requires the densities estimation of genuine and impostor match scores [8] (e.g., existing approaches are based on use of kernel density estimator (KDE) [9] or mixtures of Gaussians (MoG) [12]). A great advantage of this class is that optimal performance can be achieved at any desired operating point, if the score densities are estimated accurately, while the main drawback is the need of a training set even larger than the previous class.

A very interesting result, published in [12,15], about the comparison of fusion methods from the three categories reports that approaches from the first class based on the simple sum rule coupled to a carefully chosen score normalization scheme (or similarly a weighted sum where the weights are related to the respective error rates of the matchers) is able to improve significantly the stand-alone systems performance, contrarily to fusion methods based on more complex trained approaches. This behavior is explained considering the difficulty of a general purpose classifier (e.g., support vector machines) to work using a very low number of features (i.e., the number of matchers employed in the fusion) [12]. On the contrary, experimental results reported in [35] reveal that SVM-based fusion could achieve better performance than sum rule-based fusion, provided that the kernel and its parameters have been carefully selected. Thus it is clear that the choice of the fusion approach is strictly problem-dependent: in this work we evaluate different fusion approaches with the aim of making a further step in the comparison of different fusion techniques in the fingerprint biometric field.

As the experimental evaluation is concerned all the fusion methods are tested on the four databases of FVC2006. FVC competitions (e.g., [1]) are well-known in the scientific and industrial community as one of the most successful attempt to establish common benchmarks for fingerprint recognition algorithms. Some previous studies on the combination of different fingerprint matching algorithms submitted to FVC2004 [16,17] or the recent FVConGoing [31] have proved the benefits and limits of the fusion of classifiers: the results of such works suggest that an integration of several fingerprint matchers offers performance gains that may not be possible with a single matcher, mainly when difficult images are tested.

In this work we integrate those results by considering some new statistical rules for combining matching scores and exploring the correlation among the matchers to understand the behavior of the fusion rules. All the tested methods are evaluated on the four datasets of FVC2006: as expected the fusion of different state-of-the-art fingerprint matchers can achieve performance which is not achievable using a single matcher.

The following machine learning approaches are tested: classifier-based score fusion; density-based score fusion. Moreover, we have also tested statistical rules that are not based on learning: sum rule, mean rule, product rule, max rule, and min rule.

Almost all the methods use as features, for describing a given fingerprint match, the scores of the competitors of FVC2006. Instead, in the proposed Likelihood approach (see Section 2.3) the features for training the classifiers are obtained using the likelihood ratio test (see Section 2.3).

The manuscript is organized as follows: in Section 2 the fusion algorithms proposed and tested in this work are briefly reviewed, in Section 3 the FVC2006 competition, including participants, databases, testing protocols and results are described, in Section 4 we discuss our experimental results and, finally, we draw our conclusions in Section 5.

2. Fusion algorithms

In this paper we aim at answering the following questions: which fusion strategy at score level can bring the best results in terms of performance and how much improvement can we expect from a combined system compared to a stand-alone one? To this aim, we examine different fusion strategies belonging to the following three categories and we perform experiments for comparison.

2.1. Transformation-based score fusion

This class includes the most simple fusion methods, based on score normalization to a common domain and combination