Pattern Recognition 47 (2014) 128–140

Contents lists available at ScienceDirect

Pattern Recognition

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

Legendre polynomials based feature extraction for online signature verification. Consistency analysis of feature combinations

Marianela Parodi *, Juan C. Gómez

Laboratory for System Dynamics and Signal Processing, FCEIA, Universidad Nacional de Rosario, CIFASIS, CONICET, Argentina

ARTICLE INFO

Available online 16 July 2013

Keywords:
Online signature verification
Legendre polynomials
Consistency factor

ABSTRACT

In this paper, feature combinations associated with the most commonly used time functions related to the signing process are analyzed, in order to provide some insight on their actual discriminative power for online signature verification. A consistency factor is defined to quantify the discriminative power of these different feature combinations. A fixed-length representation of the time functions associated with the signatures, based on Legendre polynomials series expansions, is proposed. The expansion coefficients in these series are used as features to model the signatures. Two different signature styles, namely, Western and Chinese, from a publicly available Signature Database are considered to evaluate the performance of the verification system. Two state-of-the-art classifiers, namely, Support Vector Machines and Random Forests are used in the verification experiments. Error rates comparable to the ones reported over the same signature datasets in a recent Signature Verification Competition, show the potential of the proposed approach. The experimental results, also show that there is a good correlation between the consistency factor and the verification errors, suggesting that consistency values could be used to select the optimal feature combination.

1. Introduction

Automatic signature verification has long been considered an important research area in the field of biometrics [1–4]. Signature verification is the most popular method for identity verification. Signatures are recognized as a legal means of verifying an individual's identity by financial and administrative institutions. In addition, it is a non-invasive biometric technique, and people are familiar with the use of signatures for identity verification in their everyday life.

Two categories of signature verification systems can be distinguished taking into account the acquisition device, namely, offline and online systems. For offline verification systems, only the image of the signature is available, while for online systems, dynamic information acquired during the signing process, such as x and y pen coordinates and pen pressure, is available. The interest in the online approach for signature verification has increased in recent years due to the widespread use of electronic pen-input devices, such as digitizer tablets and PDAs. In addition, it would be reasonable to expect that the incorporation of dynamic information acquired during the signing process would make signatures more difficult to forge and, in this way, the online verification systems more reliable than the offline ones.

In online systems, the signature is parameterized by several discrete time functions, *e.g.*, pen coordinates, pen pressure and, when available, pen inclination angles. Researchers have long argued about the effectiveness of these different time functions for verification purposes. During the First International Signature Verification Competition (SVC2004), the results using only pen coordinates outperformed those adding pen pressure and pen inclination angles [5]. Since then, several works have been presented concerning the best set of features to model the signatures. In [6], the authors state that using only pen coordinates leads to better results than incorporating the pen pressure. The time variability between training and testing data acquisition sessions is considered in [7], where it is concluded that pen pressure is the most unreliable feature, pen inclination angles are too unstable, and pen coordinates are the most robust time functions in the presence of a long term time variability. On the other hand, some works show improvements when combining pen coordinates with pen pressure and inclination angles [8]. The conflicting results observed in the literature make the discussion still open. In a preliminary work by the present authors [9], some feature combinations based on the pen coordinates and the pen pressure, are studied. The conclusions in [9] are in line with the idea that combining pen coordinates with the pen pressure leads to a verification performance improvement.

* Correspondence to: CIFASIS, 27 de Febrero 210 bis, (S2000EZP) Rosario, Argentina. Tel.: +54 341 4237248x335; fax: +54 341 4821772x3.
E-mail addresses: parodi@cifasis-conicet.gov.ar (M. Parodi), gomez@cifasis-conicet.gov.ar (J.C. Gómez).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.06.026

A desirable property for any feature is to have high consistency in the sense that the feature values of the genuine signatures should be close to each other while the ones of genuine and forged signatures should be not. A well defined consistency model would allow to quantify the discriminative power of the features and to predict their effectiveness for verification purposes. A consistency model was first introduced in [10,11]. In [10], the consistency model is used to select an optimal subset of global features from a larger global feature set. In [11], several local and global features are compared on the basis of their consistency, resulting pen coordinates and some derived features the most reliable ones. The lack of a widely used consistency model in the literature, makes its study an interesting issue. In [9], a new consistency factor is introduced. The proposed feature combinations are compared based on their consistency factor values, being the feature combinations containing the pen pressure the most reliable ones.

An important factor that deserves more investigation is the influence of the cultural origin of the signatures in the performance of the verification systems. To the best of the authors' knowledge, there are not many works in the literature that consider non-Western signatures such as Chinese, Japanese, Arabic, etc. In [12], an updated survey of non-English and non-Latin signature verification systems can be found. Non-Western signatures do have different shapes and the writing style is different to the Western one. For instance, the Chinese handwriting style consists of one or more multi-trace characters, most of them being phono-semantic compounds, composed by two parts: the radical, which is often a simplified pictograph and suggests the character's general meaning and a phonetic indicator. Originally, Chinese pictographs conveyed their meaning through pictorial resemblance to a physical object. Although in modern Chinese this resemblance is no longer clear, Chinese characters are still pictorial symbols. Among the literature of non-Western signature verification, more attention has been given to Chinese signatures than to Japanese, Arabic, Persian or Indian signatures. Offline [13,14], as well as online [15] verification systems have been presented in the literature for Chinese signature verification. Further, the Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011) held within ICDAR 2011 [16], introduced a new Database containing Chinese signatures, encouraging the researches to work on this type of data. On the other hand, Japanese and Arabic signatures, among others, have not been investigated so much. Japanese signatures consist of different component characters spaced from each other. There is not much work done on this type of data [17,18], and it is mostly focused in offline data. Arabic script is written from right to left in a cursive style. Although a lot of research has been carried out on Arabic handwriting recognition, not much work has been carried out on Arabic signature verification. In [19], an offline verification system for Arabic signatures is presented. For a verification system to have a widespread acceptance it should take into account these different writing styles. As pointed out in [12], there are still many challenges in this research area.

In this paper, the coefficients in the Legendre series approximations of the time functions associated with the signatures are used as features to model them. The time functions considered in this paper are the pen coordinates, pressure, velocities and acceleration, as well as the log curvature radius, which are the most commonly used functions in the literature for online signature verification [20,21]. A consistency factor is proposed to quantify the discriminative power of different combinations of the time functions related to the signing process. Two different signature styles are considered, namely, Western and Chinese, of a publicly available Signature Database. Two state-of-the-art classifiers, namely, Support Vector Machines (SVMs) and Random Forests (RFs), are used to perform the verification experiments.

This approach of representing the time functions using Legendre polynomials was first introduced by the present authors in the conference paper [9]. Only few feature combinations were studied there, and a qualitative study of the consistency of the feature combinations was performed. In the present paper, more time functions are considered and a thorough analysis of all the possible feature combinations is carried out. In addition, a quantitative study correlating the consistency factor with the verification errors is performed.

The main contributions of this paper are the following:

- A feature extraction approach based on Legendre series representation of the time functions associated with the signatures is proposed. To the best of the authors' knowledge this is the first time that this approach is used in the context of signature verification.
- A consistency factor is proposed to quantify the discriminative power of different combinations of the time functions associated with the signing process. A thorough study of all the possible feature combinations is carried out, and the pros and cons of these different combinations are analyzed.
- A quantitative study of the relationship between the proposed consistency factor and the verification performance of a feature combination is performed based on correlation analysis.
- The experiments are performed on one of the most recent signature datasets, containing Western and Chinese signatures, which have been used in one of the latest signature verification competitions. To quantify the verification performance, the EER (Equal Error Rate) and the cost of the log-likelihood ratios $\hat{C}_{llr}$ are reported.

The paper is organized as follows. The feature extraction approach is described in Section 2. In Section 3 the proposed consistency model is introduced. In Section 4 the Database is described. Section 5 is devoted to the description of the experiments, in particular, Sections 5.1 and 5.2 focus on the consistency computation and the verification experiments, respectively. In Section 6 the results are presented and discussed. Finally, some concluding remarks are given in Section 7.

2. Feature extraction

Several methods have been proposed in the literature for online signature verification. They differ basically in the way they perform the feature extraction and in the classification approach they employ. The different features can be classified into local features, calculated for each point in the time sequence, and global features, calculated from the whole signature. Many researchers accept that approaches based on local features achieve better performance than the ones based on global features, but still there are others who favor the use of global features [21,22]. When using global features, feature vectors have a fixed amount of components regardless the signature length. This represents an advantage since it makes the comparison between two signatures easier with respect to the case of having different feature vector lengths. Moreover, a fixed-length model of the signatures can be required for certain biometric applications [23,24]. In [25], a fixed-length representation of the signatures is proposed based on the Fast Fourier Transform (FFT). In this paper, the approximation of the different time functions by orthogonal polynomials, introduced by the authors in [9], is employed to obtain a fixed-length representation of the signature.

2.1. Basic functions

Typically, the measured data consists of three discrete time functions: pen coordinates x and y, and pen pressure p. Depending