Efficient Off-Line Verification and Identification of Signatures by Multiclass Support Vector Machines

Emre Özgündüz, Tülin Şentürk, and M. Elif Karslıgil

Computer Engineering Department, Yıldız Technical University,
Yıldız , Istanbul, Turkey
emre_ozgunduz@yahoo.com, tulinsenturk@hotmail.com
elif@ce.yildiz.edu.tr

Abstract. In this paper we present a novel and efficient approach for off-line signature verification and identification using Support Vector Machine. The global, directional and grid features of the signatures were used. In verification, one-against-all strategy is used. The true acceptance rate is 98% and true rejection rate is 81%. As the identification of signatures represent a multi-class problem, Support Vector Machine's one-against-all and one-against-one strategies were applied and their performance were compared. Our experiments indicate that one-against-one with 97% true recognition rate performs better than one-against-all by 3%.

1 Introduction

Handwritten signature recognition is a behavioral biometric technique for personal identification. Signatures are usually composed of special characters and picture-like patterns. In contrast to the unique and stable biometric features such as fingerprint and iris, even the sequentially signed signatures of a person can vary. Nevertheless as signatures are the primary mechanism both for authentication and authorization in legal transactions, the need of efficient automated solutions for signatute recognition is important.

In biometric applications, there are two types of identity recognition methods: verification (authentication) and identification. For the signature recognition, verificiation is the decision about whether the signature is genuine or forgery and identification is finding the owner of the signature. In the decision phase, the forgery images can be classified as random, simple and skilled [1]. In random forgeries, the signatures are signed without knowledge about the name and genuine signature of the owner. Simple forgeries define the signatures where the name of the signature owner is known and finally in skilled forgeries the aim is to make an almost exact copy of the genuine signature by using an existing sample.

Signature Recognition systems are categorized as on-line and off-line systems according to their applications. In the off-line systems, signature images are obtained by a scanner and usually shape characteristics are examined for the recognition. In the online systems data are obtained by a digitizing tablet. In addition to shape of the signature dynamic features as speed, stroke, pen pressure and signing duration are also analyzed.

A. Gagalowicz and W. Philips (Eds.): CAIP 2005, LNCS 3691, pp. 799 – 805, 2005.

There are several implementations for signature recognition and verification. Justino, Bortolozzi and Sabourin proposed an off-line signature verification system using Hidden Markov Model [2]. Zhang, Fu and Yan (1998) proposed handwritten signature verification system based on Neural 'Gas' based Vector Quantization [3]. Vélez, Sánchez and Moreno proposed robust off-line signature verification system using compression networks and positional cuttings [4]. Arif and Vincent (2003) concerned data fusion and its methods for an off-line signature verification problem which are Dempster-Shafer evidence theory, Possibility theory and Borda count method [5]. Chalechale and Mertins used line segment distribution of sketches for Persian signature recognition [6]. Sansone and Vento (2000) increased performance of signature verification system by a serial three stage multi-expert system [1].

In this paper a novel approach to off-line signature verification and identification using Support Vector Machine(SVM) is proposed. Support Vector Machine is a new learning method introduced by V.Vapnik and his co-workers [7] [8]. With a set of training examples belonging to two classes, Support Vector Machines finds the optimal separating hyperplane, which maximizes the minimum distance from either class to the hyperplane. Therefore the misclassification error of unseen data is minimized. The training points on the border are support vectors. Even with many features present, only support vectors are used for classification. In practice, the data may not be linearly separable. In this case, data map into a higher dimensional feature space by a kernel function and construct an optimal hyperplane in this space [9]. The commonly used kernel functions are polynomial, radial basis, and sigmoidal.

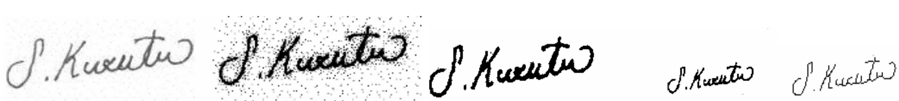
2 Preprocessing

The operations in preprocessing phase make signatures normalized and ready for feature extraction. The preprocessing steps are binarization, noise reduction, width normalization and skeletonization. An example of preprocessing is shown in Fig. 2.

Binarization: The signatures are scanned in gray level. To separate signatures from background, p-tile thresholding is applied. The pixels belonging to signature are changed to black and background is changed to white.

Noise Reduction: The small noises in the image is eliminated by a simple noise reduction filter. For each black pixel, if the the number of neighbors in white is more than number of neighbors in black, the pixel color is changed to white.

Width Normalization: Signature size may have interpersonal and intrapersonal differences. To compare two signatures their length must be the same. To provide this, the signature width is set to a default value and the height is changed with respect to height-to-width ratio[10].



(a) (b) (c) (d) (e)

Fig. 2. Preprocessing steps: (a) scanning, (b) background elimination, (c) noise reduction, (d) width normalization, (e) skeletonization applied signature