Ebrahim, AY, Kolivand, H, Rehman, A, Rahim, MSM and Saba, T

Features selection for offline handwritten signature verification: State of the art

http://researchonline.ljmu.ac.uk/id/eprint/9857/

Citation (please note it is advisable to refer to the publisher’s version if you intend to cite from this work)


LJMU has developed LJMU Research Online for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk
Features selection for offline handwritten signature verification: Current state of the art.

AY Ebrahim, H Kolivand, A Rehman, MSM Rahim, T Saba

Abstract

This research comes out with an in-depth review of widely used techniques to handwritten signature verification based, feature selection techniques. This paper focuses on selected best features of signature verification, characterized by the number of features representing for each signature, where the objective is to discriminate if a given signature is genuine, or a forgery. We present how the discussion of the advantages and drawbacks of each from feature selection techniques, has been handled by several researchers in the past few decades and the recent advancements in the field. Experiments were conducted using databases for signature verification systems (GPDS). Results were tested using two standard protocols; GPDS and the program for rate estimation and feature selection. The current precision of the signature verification techniques reported in state of art is compared with benchmark database and possible solutions are suggested to improve the accuracy. As the equal error rate is an important factor for evaluating the signature verification’s accuracy, the results show that the feature selection methods has successfully contributed toward efficient signature verification.

Keywords: Handwritten Signature Verification; Feature Extraction; Feature Reduction Methods; Feature Selection.

1 Introduction

The identification of individuals, handwritten signature is widely used and accepted mechanism throughout the world, the thorough scrutiny of the signature is important before going to the conclusion about the signer. This variance in genuine signature makes it difficult to differentiate between the genuine and forged signature. The automated Signature verification system may improve the authentication process and can differentiate between the original and forgery signatures [1]. The handwritten signature has also an adequate importance in online banking applications, credit cards, and cheque processing mechanism [2]. For the authentication and validation of passports, biometrics systems can be used; specifically for the signature verification [3].

Features extraction can be defined as the characteristics of signature that are derived from that signature itself. These extracted features play an important role in developing the robust system as all other phases are based on these features Based on, know that a large number of features may decrease the value of FRR (overall amount of genuine signatures discarded by the system) but at the same time it will increase the value of FAR (number of forged signatures accepted by the system). However, little work has been done in measuring the consistency of these features. This consistency measurement is important to determine the effectiveness of the system. In order to measure the consistencies of these features, there is a need to select the best features set among them [4][5]. There are two main tasks of signature recognition and verification one of them is the correct identification of the owner of the signature, and the other is a correct classification of signature whether it is a genuine or a
forged [6]. The handwritten signatures are the most authentic and realistic use of a person’s identification in legal and commercial transactions [7].

The focus of this paper will be on offline signature verification techniques. Further division of this paper will be, in Section II contains the literature review of the already published existing techniques of offline signature verification, Section III contains the critical analysis table of some research papers, and finally, in section IV, the conclusion of paper will be presented.

2 Research background

In the literature of Offline Signature Verification, we can find multiple ways of defining the problem. In particular, one matter is critical to be able to compare related work: whether or not skilled forgeries are used for training. Some authors do not use skilled forgeries at all for training (e.g. [8]), other researchers use skilled forgeries for training writer-independent classifiers, testing these classifiers in a separate set of users (e.g. [9]); lastly, some papers use skilled forgeries for training writer-dependent classifiers, and test these classifiers in a separate set of genuine signatures and forgeries from the same set of users.

Boosting feature selection is performed by features selection methods that select the single most discriminant feature of a set of potential features and finds a threshold to separate the two classes to learn, effectively a decision stump. Consequently, features are selected in a greedy fashion according to the weighting while learning is conducted by the features selection methods. Given a very large set of features, the result is a committee built on the best subset of features representing the training data [13]. Feature selection techniques are important also for detection of breast cancer by enhance the appearance of the mammogram images and highlight suspicious areas. Also it extracted certain dynamic features to best distinguish between benign and malignant mammograms [10], diagnosing of breast cancer based on feature selection by the Medio lateral oblique fragment of the pectoral muscle. The of breast region is performed using the multilevel wavelet decomposition of mammogram images. [11]. In addition, the image enhancement is followed by player and face detection, face recognition based on feature selection. The algorithm based on multi-scale retinex is proposed for image enhancement. Then, to detect players' and faces', used adaptive boosting and Haar features for feature extraction and classification [12]. The concept of feature selection proposes a system for signatory recognition which is based on reducing the number of features from the signature [14]. Proposed a good approach to feature selection, which when applied for signature provides a good way of compressing the signature while maintaining acceptable identification rates.
3 Signature Verification

Handwritten signatures have applied as biometric features that distinguish persons. It has confirmed that signature samples are a very faultless biometric feature with a low conflict percentage. Some signature samples might be similar but there are different scientific mechanisms to distinguish between them and for disclosure of forged signatures. There are two types of Signature Verification Systems:

3.1 Verification System of Offline (Static) Signature

Signature is written offline like a signature written on bank checks and the technique read the scan image of the signature then obtains it with the signature samples stored in the database. Off-line signatures are shown in Figure 1 [15].

![Offline signatures](image)

**Figure 1** Offline signatures

3.2 Verification System of Online (Dynamic)

Signature signing onto a reactive electronic system for example in a tablet and is read online, and comparison of signatures on file of the individual to test for authenticity. Several best features are used with online signature samples that are not accessible for the offline ones. Online signature is displayed in Figure 2 [15].
4 Benchmark datasets

The availability of datasets is one of the most important requirements in any research area. So the same is the case with signature analysis and recognition. A number of datasets comprising signature samples have been developed over the years mainly to support signature verification, signature segmentation, and signer recognition tasks. Especially, during the last few years, a number of standard datasets in different scripts and languages have been developed allowing researchers to evaluate their systems on the same databases for meaningful comparisons. Some notable dataset of signature samples along with their exciting measurements are defined in Table 1.

<table>
<thead>
<tr>
<th>Dataset Name</th>
<th>Language signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPDS [16]</td>
<td>English 8640</td>
</tr>
<tr>
<td>CEDAR [17]</td>
<td>English 2640</td>
</tr>
<tr>
<td>Arabic dataset [18]</td>
<td>Arabic 330</td>
</tr>
</tbody>
</table>
5. Preprocessing

For effective recognition of a signatory from offline signature samples, the signature must be distinguishable from the background allowing proper segmentation of the two. Most of the signatory identification techniques developed to date rely on features which are extracted from binary images with white background and black ink trace. An exception to this is the work of Wirotius et al. [21], where the authors argue that like an online signature sample, grayscale images also contain information about pen pressure, the intensity of the gray value at a particular pixel being proportional to the pen pressure.

Zuo et al. [22] also supported this idea and conducted a series of signatory identification experiments both on grayscale and binary images. The experiments on grayscale images reported slightly better results than the binary images with an overall identification rate of 98%. It should, however, be noted that feature extraction from the gray ink trace is quite complex as opposed to the binary version. A large number of useful features can be extracted from a binarized version of signature and consequently, most of the contributions to signatory identification are based on binary images of signature [22]. A number of standard thresholding algorithms have been developed to binarize images into foreground and background [23], and these methods can also be applied to signature samples. Most of the research employs the well-known Otsu’s thresholding algorithm [23], to compute a global threshold for the signature image and convert the gray scale image into binary [24].

Signature images may present variations in terms of pen thickness, scale, rotation, etc., even among authentic signatures of a person. Common preprocessing techniques are signature extraction, noise removal, application of morphological operators, size normalization, centering and binarization. The experiments on grayscale images reported slightly better results than the binary images with an overall identification rate of 98%. It should, however, be noted that feature extraction from the gray ink trace is quite complex as opposed to the binary version. A large number of useful features can be extracted from a binarized version of signature and consequently, most of the contributions to signer identification are based on binary images of signature [24].

6 Feature Extraction

This phase represents each writing fragments, by types of features. this features gives enough information about strokes in each fragment, but may be other features will not helpful in this study. These features include Local and global features.

6.1 Global and local feature extraction

Local and global features include data, which are efficient for signature confession. The features choosing is various features is vital for any style confession and classification method. Global attributes are extracted from the complete signature.
The set of these local and global attributes is further applied for reporting the identity of documentation and forgery signature samples from the dataset. The global attributes that are extracted from the sample are described as follows [25].

- **Width (Length)**: For a binary signature, width is the dimension of 2 pixels in the horizontal projection and must include more than three points of the signature.
- **Height**: Height is the dimension between two pixels in the vertical projection and must include more than three pixels of the image for a binary image.
- **Aspect ratio**: The Ratio is a global attribute that represents the ratio of the width and the height of the image [26].
- **Horizontal projection**: Horizontal projection is calculated from both binary and the skeletonized signatures. The number of black pixels is calculated from the horizontal projections of binary and skeletonized images.
- **Vertical projection**: A vertical projection is defined as the number of black pixels achieved from vertical projections of binary and skeletonized images.
- **Upper and lower edge limits**: The variation between smoothened and original curves of vertical projection is known as lower and upper edge limits, respectively.

Local attributes extracted from gray level, binary signatures. From the small regions of the whole image, local features represent, height, width, horizontal, aspect ratio and vertical projections etc. To get a group of global and local attributes, both of these feature groups are collected into a feature vector and the feature vector is represented as input to the classifiers for generating match [27].

### 6.2 Orientation

Orientation defines the direction of the signature lines. This feature is necessary due to it lets us know how the signatory signed down the signature, which letters came first emphasizing towards angles and peaks. Orientation is obtained by using the proportion of angle a major axis [28].

### 6.3 Eccentricity

Eccentricity is defined as the central point in an object. In the instance of signature, eccentricity is the central point of the image. The significance of this feature is that we want to know the central point of two signatures in order to compare them. The central point is obtained by using the proportion of the major to the minor axis of a signature [29]. In the last decades, offline signature verification has been studied which extracts the features of the signature depending on the method used. In offline method local and global features such as aspect ratio that is the ratio of height to width, grid, orientation, eccentricity and contour features are studied to get the signature in terms of its features [30]. Local attributes extracted
from gray level, binary signatures. From the small regions of the whole image, local features represent, height, width, horizontal, aspect ratio and vertical projections etc. To get a group of global and local attributes, both of these feature groups are collected into a feature vector and the feature vector is represented as input to the classifiers for generating match [28].

7. Dimension Reduction Methods

Overall, this research addresses, dimension reduction, problems in classification for high-dimensional multivariate. Figure 3 illustrates the basic idea of this study schematically. Dimension reduction has to associate with the final classification, i.e. when reducing the dimension of the data space. Figure 3, represents how the work of the data reduction methods. Classification problems refer to the assignment of some combination of input variables, which are measured or preset, into predefined classes. Over the last decade, technological advances have brought the large growth of data dimension, where the number of variables is often in the hundreds of thousands and considerably larger than the number of observations. This problem has influenced a broad range of areas such as image processing and text data analysis.

![Figure 3 Representation data of the data reduction methods](image)

This research provides an overview of dimension reduction for multivariate data. Generally, the aim is to reduce the dimension of X without loss of data in Y|X, and without requiring a specific model for Y|X, where X is the predictor and Y is the response. There are three advantages of dimension reduction, it decreases the time, storage and space required. The first part of dimension reduction is feature selection approaches, which is a try to find a subset of the original variables (also called features or attributes). In some cases, data analysis such as regression or classification can be done in the reduced space more accurately than in
the original space such as Sparse PCA technique which is discussed in detail in the next subsection [31]. In recent years a variety of linear and nonlinear reduction technique has been proposed many of which rely on the evaluation of local properties of the data. This research presents a review and systematic comparison of these techniques. By identifying the weaknesses of current, linear and nonlinear techniques.

7.1 Linear Dimension Reduction

Linear techniques perform dimension reduction by embedding the data into a subspace of lower dimension. There are various techniques to do so, such as linear discriminant analysis (LDA) and principal component analysis (PCA), [31]. Linear Discriminant Analysis (LDA) is a popular data-analytic tool for studying the class relationship between data points and LDA is supervised for searches for the project axes on which the data points of different classes are far from each other while requiring data points of the same class to be close to each other. A major disadvantage of LDA is that it fails to discover the local geometrical structure of the data manifold [32].

Dimension reduction is the task of reducing the amount of available data (a.k.a. data dimension). The data processing required in dimension reduction, ten times linear for computational simplicity, is determined by optimizing an appropriate figure of merit, which quantifies the amount of information preserved after a certain reduction in the data dimension. The ‘workhorse’ of dimension reduction comes under the name of principal component analysis (PCA), [31]. PCA has been extremely popular in data dimension reduction since it entails linear data processing. PCA is by far the greatest popular linear technique. Therefore, in comparison only include PCA as a benchmark [33]. PCA which will discuss detail in next subsections.

7.2 Non-Linear Techniques for Dimension Reduction

Most nonlinear techniques for dimension reduction have been proposed more recently and are therefore less well studied. This section will discuss two nonlinear technique for dimension reduction which is called (i) Kernel PCA (KPCA) and (ii) Multi-dimensional scaling (MDS). These are discussed techniques that attempt to preserve global properties of the original data in the low-dimensional representation.

KPCA is the reformulation of traditional linear PCA, shown in Figure 4 KPCA computes the kernel matrix $K$ of the data points $x_i$. Kernel PCA is a Kernel-based method. As shown in Figure 4 the mapping performed by Kernel principal component, are lies on the choice of the kernel function. An important weakness of Kernel PCA is that the size of the kernel matrix is proportional to the square of the number of instances in the dataset, and the problems in finding the smallest eigenvalues in an Eigenproblem [33].
Multi-Dimensional Scaling (MDS) is a mathematical dimension reduction technique that maps the distances between observations from the original (high) dimensional space into a lower dimensional space. Honarkhah et al. [34] represents a collection of nonlinear techniques that map the high dimensional data representation to a low-dimensional representation while retaining the pairwise distances between the data points as much as possible. The summon of cost function differs from the raw stress function in that it puts more emphasis on retaining distances that were originally small. A major disadvantage Multi-Dimensional Scaling (MDS) provides a global measure of dis/similarity but does not provide much insight into subtleties. The susceptibility to the curse of dimension and the problem in finding the small eigenvalue in an eigenproblem.

From these observations, it is clear that nonlinear techniques impose considerable demands on computational resources, as compared to the linear technique. The second disadvantage of the PCA consists in the fact that the directions maximizing variance do not always maximize information. In this case, the PCA will give preference to the first (less informative) variable. This drawback is closely connected to the fact that the PCA does not perform linear separation of classes, linear regression or other similar operations, but it merely permits the input vector to be best restored on the basis of the partial information about it. All additional information pertaining to the vector (such as the identification of an image with one of the classes) is ignored. PCA is sensitive to the relative scaling of the original variables [35]. Feature extraction creates new features from functions of the original features, whereas feature selection returns a subset of the features. Feature extraction transforms the data in the high-dimensional space to a space of fewer dimensions [35]. The data transformation may be linear, as in principal component analysis (PCA), PCA is an unsupervised method. It aims to project the data along the direction of maximal variance.

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of PC is less than or equal to the number of original variables as shown in Figure 5.
In Figure 6 PCA can be found via compute the SVD of matrix factorization is $X^T X$ complex matrix (empirical covariance matrix of $X$). $V$ represents vector, singular value decomposition (SVD) is a factorization of a real or complex matrix.

In Figure 7 Each eigengene is expressed only in the corresponding Eigen- array with the corresponding Eigen-expression level. PCA can be found via compute the SVD of the features matrix. Compute the SVD of $X^T X B = U D V^T$, where SVD is singular value decomposition, UD are the principal components, the columns of $V^T$ are the consistent loading of the primary components eigenvectors, $V$ of which diagonalizes the covariance matrix $X^T X$. 
In Figure 7 the variance of X that is remained in X’ is maximal. Dataset X is mapped to dataset X’, here of the same dimension. The first dimension in X’ (e_1 = the first principal component) is the direction of maximal variance. The second principal component (e_2) is orthogonal to the first.

![Figure 7](image)

**Figure 7** Eigenvalue measures the variation in eigenvectors e

The main drawback of PCA is that the size of the covariance matrix is proportional to the dimensionality of the data points. As a result, the computation of the eigenvectors might be infeasible for very high-dimensional data [35]. The other problem is that PCA is sensitive to outliers because it computes eigenvalues and eigenvectors based on the conventional covariance matrix. Consequently, there might be situations where the components explain a structure created by a relatively small number of outliers.

8 Feature Selection

Features selection is the process of choose a subset of relevant features for use in techniques in the offline signature identification and signature verification. In some cases, the current feature does not improve the capability, these features are too many (high dimensions), which reduce classification process efficiency for this we need to selected best features, from features extraction as shown in Figure 9. Many researchers [36], [13] proposed features selection techniques to select features from the signature image and achieved good quality results. Many papers have used a feature selection approach for signature verification. Trained a writer-independent classifier, by first extracting a large number of features from each signature (over 30 thousand features), applying feature extractors at different scales of the image. Their method consisted in training an ensemble of decision stumps (equivalent to a decision tree with only one node), where each decision stump only used one feature. With this approach, they were able to obtain a smaller feature representation (less than a thousand features) that achieved good results in the Brazilian and GPDS datasets. Eskander et al. [36] extended Rivard’s [13] work to train a hybrid writer-independent-writer-dependent classifier,
by first training this writer-independent classifier to perform feature selection, and then train writer dependent classifiers using only the features that were selected by the first model. This strategy presented good results when a certain number of samples per user is reached.

Feature selection methods applied for three causes: facilitation of methods to make them easier to explain by writers, build better, faster, and easier to understand learning machines and shorter training times. Finally, enhanced generalization by reducing of variance.

Feature selection techniques are often used in domains where there are many features and comparatively few samples for the implementation of feature chosen contain the analysis of signatures where there are many thousands of features and a few tens to hundreds of images. Features are selected to add while building the model based on the prediction errors. In some situations, information analysis such regression or classification can be done in the reduced area more accurately than in the main area. In this study, three techniques, Principal Components Analysis (PCA) and Sparse Principal Components Analysis (SPCA) are used for feature selection and reduction. As has been shown in Figure 9, since features dimensions of features selection are large with redundant features, reducing of their dimension is must to choose the most important ones. In the presence of many features, select the most relevant subset of combinations of features.

![Figure 9](image_url) Representation best features of the features selection methods as input to classification technique.
The features which are extracted for signature, such as Kalera et al. [17] who extracted 3 types of features, Biswas et al. [37] extracted 5 types of features and Pushpalatha et al. [38], Pouraza et al, [27] extracted 8 types of features in offline signature verification. Siddiqi and Vincent [39] extracted 10 type of features in Offline Handwritten Recognition. A.Y. Ebrahim [5] extracted 20 types of features in offline signature verification as shown in Table 2.

**Table 2:** Number of Features Extraction Used in the Classification, Identification and Verification Process

<table>
<thead>
<tr>
<th>Author/year</th>
<th>Number of Features</th>
<th>Types of Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalera (2004) [17]</td>
<td>3 types of Features extracted in Offline Signature Verification</td>
<td>Eccentricity, rectangularity, and orientation</td>
</tr>
<tr>
<td>Daramola (2010) [26] and Pouraza (2011) [27]</td>
<td>8 types of Features extracted in Offline Signature Verification</td>
<td>Vertical projections, horizontal projections, upper profiles, lower profiles, elongation, solidity, eccentricity, and rectangularity, Number of Cross-points, Number of edge points, eccentricity, Mass and Centre of Mass.</td>
</tr>
<tr>
<td>Pushpalatha (2014) [38]</td>
<td>5 types of Features extracted in Offline Signature Verification</td>
<td></td>
</tr>
</tbody>
</table>

comparison is presented between other existing techniques on the basis of feature extraction, feature selection, classification techniques and performance measures such as verification rate, as shown in Table 3.
Table 3: Related study on off-line signature verification techniques

<table>
<thead>
<tr>
<th>S/ N</th>
<th>Author / Year</th>
<th>Methodology</th>
<th>Database</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Pourreza et al. (2011)</td>
<td>Used edge detection, Curvelet transform, Hough transform as extraction methods.</td>
<td>Train 150 Test 60</td>
<td>Improved a method depend on Monolithic ANN and (MNN) with 83%-96.6% accuracy</td>
</tr>
<tr>
<td>2.</td>
<td>Anwar et al. (2014)</td>
<td>Based on Codebook Generation technique Using ANN</td>
<td>100 users of GPDS database</td>
<td>The features are extracted using window technique, the verification accuracy rate of 96% and ERR value of 7.23</td>
</tr>
<tr>
<td>3.</td>
<td>Kaur et al. (2015)</td>
<td>Based on Surf Features Using Hmm</td>
<td>50 users of Punjabi database</td>
<td>The features are selected using SURF features and critical point matching shows the verification accuracy rate of 97%.</td>
</tr>
<tr>
<td>4.</td>
<td>Hafemann (2016)</td>
<td>Writer-independent Feature Learning for Offline Signature Verification using Deep Convolutional Neural Networks</td>
<td>160 GPDS database</td>
<td>It is better to learn the features from data shows ERR value of 10.70</td>
</tr>
<tr>
<td>5.</td>
<td>A.Y. brahim (2017)</td>
<td>DCT+WDT Features Technique</td>
<td>Arabic Signature</td>
<td>The features are selected shows the verification accuracy rate of 99.75%.</td>
</tr>
</tbody>
</table>

10 Conclusion

This study summarizes some literature reviews related to the signature verification domain, included bio-metrics for automatic verification of a person, datasets is representing signature sample of individual. In spite of these advancements, the experimental results still report somewhat large error rates for distinguishing genuine signatures and skilled forgeries, when large public datasets are used for testing, such as GPDS. This paper presents a practical solution to some of the fundamental problems encountered in the design of off-line signature
verification, the limited number of users and, the large number of features from signatures, the high intrapersonal variability of the signatures, and the lack of forgeries as counterexamples. A new approach for feature selection is proposed for the accurate design of off-line signature verification systems. It combines feature extraction, feature selection. Recently, feature selection methods with classification techniques based signer verification have emerged as an effective method for characterizing the signer of a signature and the results of these methods are found to be better than other features for signature identification and verification. As a conclusion, the method of selecting the best features among huge features will help to improve the performance of signature verification. As this paper contains the review of literature in continuation to this the next objective will be to propose some new model that will reduce the FAR and FRR.

ACKNOWLEDGEMENT
The writers are thankful to the Babylon University, Babylon, Iraq for providing study means in achieves this project.

Conflict of interest
Authors declare that they have no conflict of interest.

References:


[40] Anwar Yahy Ebrahim, Ghazali Sulong. (2014). Offline Handwritten Signature Verification Using Back Propagation Artificial Neural Network Matching Technique. 31st July. Vol. 65 No.3 JATIT & LLS. All rights reserved.
