A Survey on Handwritten Signature Verification Approaches

Basheer Mohamad Al-Maqaleh  
Faculty of Computer Science & Information Systems, Department of Information Technology, Thamar University, Thamar, Republic of Yemen

Abdulbaset Mohammed Qaid Musleh  
Faculty of Computer Science & Information Systems, Department of Computer Science, Thamar University, Thamar, Republic of Yemen

ABSTRACT
The most common secure personal authentication in biometries is handwritten signature. Signature authentication is the process of verifying an individual’s identity. Signature verification system can be classified into two methods: online and offline. This paper explains the survey of various approaches related to offline signature verification systems. Further, this paper provides the comparison of those approaches, their results and techniques of feature extraction.

Keywords
Forgery, Global Features, Local Features, Signature Verification.

1. INTRODUCTION
Biometric systems are used to identify the person through the physiological, psychological and behavioral characteristics [1]. For example, behavioral characteristics include (Voice, Signature/handwriting, Keystroke dynamics, Gait and Heartbeat, etc.) Psychological characteristics include (fingerprint, iris, face and hand geometry, etc.). Biometrics are primarily used for verification, identification, and authentication. Handwritten signatures are one of the most widely used behavioral biometries for personal identification and verification. Even with the introduction of new technologies, handwritten signatures are continuously used as a means of communication in day-to-day life like, in a formal agreements, financial systems, government use, marketing documents or paintings, etc. The main difficulty observed in a signature verification is the inconsistency of individual’s signature: variation may appear due to signing position, pen weight, width, stress, mood, time, etc. [2]. Signature verification is divided into offline and online systems.

Offline signature samples, which can be attainable by scanners or digital cameras [3]. Online signature samples are directly collected from a digitizing tablet which is capable of pen movements recording during the writing. The set of dynamic information like speed and pressure are captured, in addition to a static image of signature features [4],[5].

Generally, three types of forgery occur in offline signatures verification [6]:

Random forgeries occur when the signer just knows the name of the person whose signature is to be signed. The random sample is taken from the genuine signature that belongs to another parson.

Simple forgery occurs when the forger person knows the person’s signature shape, but has not practiced much on it.

Skilled forgery is a reasonable imitation of the genuine signature model [6]. These three different types of forgeries are shown in Figure 1.

![Types of Forgeries](image)

Fig 1: Types of Forgeries (a) Genuine Signature (b) Random Forgery (c) Simple Forgery (d) Skilled Forgery.

2. SIGNATURES DATASETS
In this section, the datasets for training and testing of the offline signature verification system are described.

- **GPDS:**
  GPDS is a database, which consists of static signature images[7],[8]. This dataset contains 24 genuine signatures and 30 simulated forgeries from 4000 individuals. All image signatures in the database are collected from a set of individuals. Each individual was singing on a sheet of white A4 paper by ballpoint. Each sheet provided two different box sizes for the signature. The images signatures in the database were scanned at 600 dpi with 256 gray levels. The training set consists of 10 randomly selected genuine signatures. The remaining genuine signatures are used as testing samples[9],[10],[11]. The database (GPDS-960) was used in [12].

- **CEDAR:**
  CEDAR dataset consists of 55 signature sets, with each set being composed of by one writer. Each writer provided 24 samples of their signature, where these samples constitute the genuine portion of the dataset. The forgeries for this dataset were obtained by asking arbitrary people to skillfully forge the signatures of the previously mentioned writers. In this fashion, 24 forgery samples were collected per writer from about 20 skillful forgers. In total, this dataset contains 2,640 signatures, built from 1,320 genuine signatures and 1,320 forgeries[1],[10].
• MCYT: The corpus MCYT is database collected in “Autonomous University of Madrid” from 75 various individuals, each person was asked to write 30 signatures (15 genuine samples and 15 forgeries). The signature images were collected to A4 size sheet of paper and was scanned by the resolution of 600dpi [13],[14],[15].

• Database special with some researchers:

  Some researcher did not used benchmarks database, but each researcher collected dataset by himself. For example in [5] the dataset consisting of 700 genuine signature images, are collected from 100 students in “Covenant University Ota Nigeria”, each student was taken 7 signatures, and 200 forgeries are collected from 20 persons, it is the mode of 100 simple forgeries and 100 skilled forgeries. In [16] the dataset which are used consist of 30 genuine signatures taken from each individual, 9 genuine signatures were selected for training stage and 21 were used in a tested stage. The forgery signatures are 30 signatures taken from three individual and 10 from each. Each volunteer may be asked to produce one to three other people’s signatures, given photocopies of the genuine signature pages. A dataset of signature images consisting of 220 genuine signatures and 110 forgeries signatures is used in [17]. It was collected from 22 people, each person signed 10 signature by different pens (fountain pen or ball pen). A dataset collected from 60 Arabic people is used in [18]. It consists of 360-signature image.

3. METHODOLOGY OF SIGNATURE VERIFICATION

The main steps used for offline signature verification system are shown in Figure 2. These steps are explained below:

![Signature Acquisition Diagram](image)

**Signature Acquisition**

**Pre-processing**

**Features extraction**

**Database**

**Signature**: The signature images are digital images of different documents using scanner or digital camera to obtain a digital image of signature[3].

**3.2 Preprocessing**

A set of the processes is necessary to preprocessing the image before feature extraction. The purpose of the signature pre-processing step is to make signatures standard and ready for feature extraction. Preprocessing is an important step to increase the accuracy of feature extraction and verification stages. These steps are shown in Figure 3 [10].

- **Binarization**: Most of methods are used for converting color image to binary image. Among these common methods are Global Threshold, Iterative Threshold, Niblack, Bernsen, Local Iterative Method and Otsu’s[19]. An Otsu’s method developed by Otsu et al., 1979, is one of the most common binarization techniques for thresholding grey-scale images [16]. It is an unsupervised clustering method for automatically choosing a global threshold that will split the signature into the foreground and background classes [9]. A second method is local iterative threshold binarization method which is used in [10].

![Fig 3: Stages of Preprocessing on an Image](image)

(a) The Original Image, (b) Binarized Image, (c) Noise-cleaned Image, (d) Rotated and Thinned Image, (e) Normalized Image.

- **Noise Removal**: Various noise filtering techniques (Like, Mean, Median, Gaussian filters, Average filter, etc.) are applied to reduce noise encounter during a scanning process[2]. Most of the common method is median noise removal. This method depends on taking the middle or median value from nine pixel's stored, as shown in Figure 4. The noise is cleaned using connected component analysis choosing a threshold $T$ ($T = 7$, is selected experimentally). Thus, isolated particles of size less than T are removed [10].

![Fig 4: Median Noise Removal](image)

- **Rotation**: Some of signature images are captured. Its rotation may be out of sync in relation to similar signatures, the image is rotated or correction rotated. Among rotation approaches are an axis of least inertia and region rotation. This approach is used in [20] to get rotation image angle, the first baseline slant angle of the image is calculated and then the image is rotated by that angle in clock-wise direction. A rotating image is not used in some of offline signature verification systems because the extracted features are not dependent on pixel location.

- **Skeletonization and Thinning**: Removing selected foreground pixels from the binary image of the signature is called Skeletonization. A representation of a signature pattern will be the outcome by a group of thin arcs and curves [21]. Eliminating the thickness difference of pen by making the image signature one pixel thick is the goal of thinning [16].

- **Normalization**: The signatures image usually may have different sizes; so, all signatures must have equal sizes to get more reasonable results through the normalization of the length or width of the image. In [18] the images are
resized to (280 x 80). The algorithm in [20] resized the acquired signatures to a standard size (100 x 200). The size (512 x 128) is used in [22] and a new formalism for signature representation based on visual perception is proposed.

### 3.3 Feature Extraction

The most challenging problem in automatic signature verification is to extract features that discriminate between genuine and forged signatures. The issue of the automated signature verification is to obtain more accurate features that enable us to differentiate between signatures. Therefore, the results of any verification system depend on the algorithm that chooses the features from the signatures. The extracted features are usually classified into the following types [23]:

- **Global features** are clearer than the other features such as width, height, aspect ratio. These features are used in combination with other features in the verification process. Knowing it’s global features are less sensitive to noise and commonly used in signature recognition process.

- **Local features** are taking after the division of the image so considered more efficient than others because they get the smallest details within the image. They are calculated by splitting the signature image into parts with the help of geometric center, density center or some other means.

- **Transition features** count the transition in the signature image from black to white pixel or vice versa in binarized signature images [24]. It used in combination with the some other features.

In [5] the suggested system partitioned the signature image into 64 cells based on the center of gravity as shown in Figure 5. The features are (F1) image cell size, image center angle relative to the cell lower right corner (F2) and pixels normalized angle relative to the lower right corner (F3). Addition to pixels normalized length relative to the lower right corner (F4) as a new feature is used in [25].

![Fig 5: Partitioned the Signature Image.](image)

A novel feature extraction method for offline signature verification is proposed by [26]. In this algorithm, six feature points are extracted based on vertical splitting and six feature points based on horizontal splitting. Figure 6 shows the signature image.

![Fig 6: Partitioned the Signature Image.](image)

An offline Arabic signature recognition and verification system designed of two phases is presented [17]. The global features and local features were used. Global features, which used are the width and the baseline. But the local features after segment signature image into 16 square based on the center-of-gravity feature are extracted, as shown in Figure 7.

![Fig 7: (a) Central lines. (b) Central line features. (c) Corner lines. (d) Corner line features. (e) Central circle. (f) Central circle features. (g) Corner curves. (h) Corner curve features. (k) Critical point features.](image)

Offline Arabic signature verification system using the combination of geometrical and grid features is presented in [18]. The geometrical features use four border points in the direction of right and left, up and down in addition to two lengths between vertical and horizontal directions. A modified offline signature verification based on geometric center features is proposed in [4]. In [1], the features used of the handwritten signatures are generated by using the curve lettrans form that is a new multi-scale transform is developed by “Candes and Donoho 1999” [27]. In [28] the features extracted are based on energy of signature on grid level. A set of features extracted in [29], are baseline slant angle, aspect ratio, normalized area, center of gravity and the slope of the line. The local parameter feature extraction method is used in [30], these features are: function features like velocity, pressure and position. In addition, parameter features are aspect ratio, closed area, grid-based approach, intersecting points, finding the contour, ratio of the distance between centroids, and ratio of the area closed by those centroids and the bounding box. In [31] a set of features was extracted to create a feature vector, these features are: horizontal and vertical histogram, center of mass, area, aspect ratio, tri surface, six fold surface and transition feature.
4. APPROACHES TO SIGNATURE VERIFICATION

Last important stage of this system is verification: This stage compares the incoming test signatures with the user’s signature template in the database. The most common approaches are:

4.1 Euclidean Distance Model

The Euclidean distance model is one of the most suitable classifier used to obtain distance measurement between two vectors of equal size on a two-dimensional plane [25]. It's used for calculating distance between extracted features as if we have a pair of vectors of size to calculate distance (d) between two vectors \(X(x_1,x_2, x_3,\ldots,x_n)\) and \(Y(y_1,y_2,y_3,\ldots, y_n)\) by using the following equation:

\[
(d) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2}
\]

(1)

Euclidean distance model is used as a classifier in [4],[5],[25],[29].

4.2 Neural Networks

A Neural Network is more efficient if we have a large number of samples. Neural Networks are highly suited to modeling global aspects of handwritten signatures. In [22] neural network used to classify extracted features from3000 signature images. The network is trained by an artificially generated genuine and forgery samples from enrollment reference signatures, which allows definite training control and at the same time significantly reduces the number of enrollment samples required to achieve a good performance. In [32], a new technique for offline signature recognition and verification is developed. The proposed technique is based on global, grid and texture features. For each one of these feature sets a special two stage Perceptron OCON (one-class-one-network) classification structure for each one of these feature sets a special has been implemented. The Euclidean distance and neural network are used in first stage as a classifier. The results of the first stage classifier feed a second stage radial base function(RBF) neural network structure, which makes the final decision. An offline verification and recognition system is presented in [33]. In this system, a multi-layer perceptron (MLP) neural network uses features like back propagation algorithm, FF algorithm, centroid method and Serialization data base for storing sample signatures which can be extracted by image processing. In addition, the neural network was trained using back propagation algorithm.

4.3 Support Vector Machines

Support Vector Machine(SVM) developed by Vapnik in (1998), and is a new technique in the field of statistical learning theory. The aim of the proposed method in [27] is to measure gray level features of an image when a complex background and train distort it by using neural network classifier and SVM. One- Class Support Vector Machine(OC-SVM) is proposed in [1] based on writer independent parameters that take into consideration only genuine signatures and when forgery signatures are lack as counter examples for designing the HSVS. The OC-SVM is effective when large samples are available for providing an accurate classification. An efficient feature extraction technique is proposed in[10] for verification using a two-class classifiers namely, RBF-SVM (support vector machine with RBF kernel) or MLP. In [34], five unique geometric features and eight Camasura features are extracted from each square. The SVM is used for the classifier.

4.4 Fuzzy Model

Various fuzzy rules are used to judge the type of signature read forged or genuine. An offline Arabic signature recognition and verification system of two phases are described in[17]. This system used fuzzy model in a verification phase for decision make once points of interest are selected, the system assigns fuzzy grades to these points depending on their degrees of matching. Offline signature verification using fuzzy Logic is proposed by [35].

4.5 Hidden Markov Model

Both offline and online signature verification approaches can used Hidden Markov Models (HMMs). Offline signature recognition using HMM is proposed by [36]. In [37]a comparison of SVM and HMM classifiers in the offline signature verification is presented.

5. COMPARISON

In an offline signature verification system, the performance is evaluated in term of error rate. The following error rates are used to compare the performance of various offline signature verification systems as shown in Table 1.

- **FRR (False Rejection Ratio):**
  
  The false rejection ratio is the ratio of the number of genuine signatures that rejected by the system to total number genuine signatures tested.

  \[
  FRR = \frac{\text{Number of originals rejected}}{\text{Number of originals tested}} \times 100 \%
  \]

- **FAR (False Acceptance Ratio):**
  
  The false acceptance ratio is given by the number of forgery signatures accepted by the system with respect to the total number of comparisons made.

  \[
  FAR = \frac{\text{Number of forgeries accepted}}{\text{Number of forgeries tested}} \times 100 \%
  \]

6. CONCLUSION

This paper presents a brief survey of various offline signature verification approaches. These approaches are studied according to their different stages, and the performance evaluation based on FRR, and FAR is given. In addition, they can analyzed for efficiency to get better result. There is a need to develop one general system in future work to classify every style of signature and to enhance performance.
Table 1. Performance comparison of various offline signature verification system.

<table>
<thead>
<tr>
<th>No.</th>
<th>Approaches</th>
<th>Features Used</th>
<th>Classifiers</th>
<th>FRR</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Novel feature extraction[5].</td>
<td>Cell size, Center angle relative to the cell lower right corner, and Pixels angle relative to the lower right corner.</td>
<td>Euclidean distance</td>
<td>0.5 %</td>
<td>1 %</td>
</tr>
<tr>
<td>2</td>
<td>Novel feature extraction[26].</td>
<td>Vertical center points, and Horizontal center points.</td>
<td>Euclidean distance</td>
<td>14.58 %</td>
<td>16.36 %</td>
</tr>
<tr>
<td>3</td>
<td>Structural feature correspondence [22].</td>
<td>Core, Outline, Ink area distribution, and Signature Frontiers.</td>
<td>Neural network</td>
<td>11.1%</td>
<td>11.8%</td>
</tr>
<tr>
<td>4</td>
<td>Globe and local features[17].</td>
<td>Central lines, Central line, Corner lines, Corner line , Central circle, Central circle, Corner curves, Corner curve, and Critical point.</td>
<td>Fuzzy</td>
<td>0.2%</td>
<td>0.2%</td>
</tr>
<tr>
<td>5</td>
<td>Combination of geometrical and grid features[18].</td>
<td>Four border points, and Lengths between vertical and horizontal directions.</td>
<td>Multi-stage classifier</td>
<td>0.16%</td>
<td>0.45%</td>
</tr>
<tr>
<td>6</td>
<td>Signature verification based on writer-independent parameters[1].</td>
<td>The energy, and Standard deviation of the curvelet coefficient.</td>
<td>SVM</td>
<td>12.50 %</td>
<td>19.40 %</td>
</tr>
<tr>
<td>7</td>
<td>An efficient feature extraction technique[10].</td>
<td>Information like connectivity among pixels, Curvilinear nature of strokes and Local density of black pixels.</td>
<td>SVM</td>
<td>13.76 %</td>
<td>13.76 %</td>
</tr>
<tr>
<td>8</td>
<td>Texture and topological features[29].</td>
<td>Baseline slant angle, Aspect ratio, Normalized area, and Center of gravity.</td>
<td>Euclidean distance</td>
<td>6.2%</td>
<td>22%</td>
</tr>
<tr>
<td>9</td>
<td>Geometric features[31].</td>
<td>Area, Eccentricity, Kurtosis, Skewness, and Centroid Co-ordinates.</td>
<td>HMM</td>
<td>2.43%</td>
<td>2.43%</td>
</tr>
<tr>
<td>10</td>
<td>Grid and centroid based approach[30].</td>
<td>Function features (like velocity, pressure, and position), Parameter features (Aspect Ratio, Closed Area, Grid-based approach, Intersecting points, Finding the Contour, Ratio of the distance between centroids, and Ratio of the area closed by those centroids and the bounding box).</td>
<td>Matching</td>
<td>7%</td>
<td>11%</td>
</tr>
<tr>
<td>11</td>
<td>Triangular geometric feature [12].</td>
<td>Sides, Angles, and Perimeter of a triangle.</td>
<td>Euclidean distance</td>
<td>2.25%</td>
<td>2.25%</td>
</tr>
<tr>
<td>12</td>
<td>Geometric feature [38].</td>
<td>Horizontal and vertical histogram, Center of mass, Area, Aspect Ratio, Tri surface feature, Six fold surface feature, and Transition feature.</td>
<td>Neural network</td>
<td>13.75%</td>
<td>13.75%</td>
</tr>
<tr>
<td>13</td>
<td>Grid and texture features [32].</td>
<td>Height, Width, Area, Baseline, Vertical center, Horizontal center, Maximum vertical and Horizontal projection, Peaks, Slant angle, Edge points, Cross points, and Closed loops.</td>
<td>Neural network &amp; Euclidean distance</td>
<td>3%</td>
<td>9.81%</td>
</tr>
<tr>
<td>14</td>
<td>Gray level features [11].</td>
<td>Gray level features.</td>
<td>SVM &amp; Neural network</td>
<td>9.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>15</td>
<td>Geometric and concentric squares approach [39].</td>
<td>Cross-points, Edge-points , Euler’s number, Mass , and Centre of mass.</td>
<td>HMM and SVM</td>
<td>4.9%</td>
<td>11.02%</td>
</tr>
<tr>
<td>16</td>
<td>An efficient offline signature verification system-using local features[24].</td>
<td>Cell size, Center angle relative to the cell lower right corner, Pixels angle relative to the lower right corner, and Pixels Length relative to the lower right corner.</td>
<td>Euclidean distance</td>
<td>10.5%</td>
<td>12.06%</td>
</tr>
</tbody>
</table>

7. REFERENCES


