**CHAPTER 1**

**INTRODUCTION**

A signature is a special case of handwriting which includes special characters and flourishes. It is a behavioral biometric which is in control of the subject and tend to change over the short and long terms due to various factors such as health, aging and physiological state. Other traits like speaking and walking also come under behavioral biometric. Physiological biometrics on other hand mainly measures the physical features of the subject such as face, figerprint, iris, hand and finger geometry. The physical features mostly remain same unlike the behavioral characteristics. A behavioral biometric generally requires several samples due to their inherent variations. Signature variations depend on fatigue, mental and physical state, and writing position.

Signature recognition and verification is one of the important ways to identify the owner of the signature and also to find whether the signature is genuine or forged. It also holds a particular importance as it the only widely accepted method for endorsing financial transactions. It has been used for decades in civilian applications while other methods still have the stigma of being associated with criminal investigation.

The signature verification can be divided into two methods which are online (dynamic) and off-line (static) verification. In this study, we are going for off-line verification where the subject’s signature is acquired on paper and scanned where biometric system recognizes the unique features and states the genuinely of it.

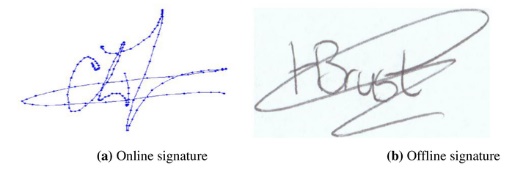
* 1. **Online Signature Recognition & Verification**

Online signature recognition and verification refers to a process where the user’s signature is acquired on a digital device. The signer uses a special pen called stylus to create their signature. Online SVRS track down path and other time-variable sequence variables using specially designed tablets or other devices during the act of signing. As Figure 1.1 shows the process where the user’s signature is acquired on a digital device.



**Figure 1.1** online signature using digital device and stylus

Automatic online SVRS is an interesting intellectual challenge with many practical applications. This technology examines the behavioral components of the signature such as: stroke order, speed, and pressure, as opposed to comparing visual images of signatures. Unlike traditional signature comparison technologies, online SVRS measures the physical activity of signing. Figure 1.2 shows the dynamic/online form of signature which is acquired on digital device by the user.



**Figure 1.2** dynamic form of signature

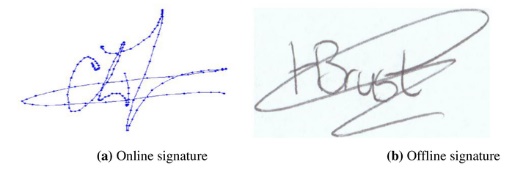
* 1. **Off-line Signature Recognition & Verification**

Off-line signature recognition and verification and recognition are concerned with the verification of signature made by a normal pen. Figure 1.3 shows the process where the user’s signature is acquired on paper using a pen.



**Figure 1.3** off-line signature using pen and paper

Off-line SVRS can be done using two different approaches. One is the writer dependent signature verification, where models for genuine and forgery signatures are constructed for each writer. Then, the test signature sample of a writer is compared to its own training sample. The second approach is as called writer-independent signature verification is used by forensic experts. This approach is considered as the most practical cases, since it is not necessary to generate a model for each writer in order to verify its signature. In this case, a general model is built from some writers chosen randomly. However, the writer-independent signature verification constitutes a more difficult task because of the important morphological variability inter-writers. Figure 1.4 shows the static/off-line form of signature which is acquired on paper using a pen by the user.



**Figure 1.4** static form of signature

**1.3 Organization of the Report**

This report starts with an overview of different ways of signature recognition and verification. We analyze the impact of different techniques and procedures to be performed on the images before comparison called preprocessing and feature extraction techniques in order to get an accurate output. The report is organized as follows:

Chapter 1: Introduction - This chapter briefly explains the overview of the report.

Chapter 2: Database - This chapter describes how the database should be collected.

Chapter 3: Preprocessing Techniques - This chapter discusses about the preprocessing techniques that are performed on the investigation image before feature extraction takes place.

Chapter 4: Feature Extraction - This chapter explains about the features to be extracted from the image in order to compare between genuine and forged signature.

Chapter 5: Feature Matching - This chapter describes the process of feature matching which happens between two images.

Chapter 6: Decision for Authentication - This chapter explains about how the output decision is made after the comparison between two images.

Chapter 7: Conclusions - This chapter summarizes the major accomplishments of this report.

**CHAPTER 2**

**LITERATURE REVIEW**

Signature verification in addition to being a popular research area in the field of pattern recognition and image processing also plays a key role in many applications such as access control, security, privacy etc. Signature verification is the task of validating a person based on his handwritten signature. There are two types of signature verification systems

1. On-line Signature Verification System which use electronic device such as a tablet to capture features like pressure speed direction etc.
2. Offline Signature Verification System in which signature is written offline and the system read the image scan then verifies it with the already stored image of the signature. Offline signature verification can be done using two different approaches.

There are many ways to recognize and verify Signatures.

K. Krishnaditya et.al [1] has proposed offline signature recognition using CNN (Convolutional neural networks). It is high proficient in areas like identification of objects faces etc. Initially signature image is pre-processed to isolate signature pixel from background noise using image processing techniques. Still have some security issues when compared with dynamic form of signature verification.

L. G. Hafemann et.al [2] has proposed signature veriﬁcation systems to discriminate if a given signature is genuine (produced by the claimed individual), or a forgery (produced by an impostor).This paper formalized the problem at hand, and list the main datasets that available to evaluate such systems. Hidden markov model was proposed for scanner resolution problem i.e. for 45 dpi to 600 dpi we need 150dpi. Feature Extraction such as Local features are extracted from a small part or a small region of the signature, Global features are extracted considering the complete signature image as a whole., and geometrical extractors measures the overall shape of a signature, Geometric features are derived from the geometrical

Parameters of the signature such as the height, width, aspect ratio, signature area etc, this includes basic descriptors, such as the signature height, width, caliber (height-to-width ration) and area. Statistical extractors some statistical features extracted from offline signatures are mean, centre of gravity of the signature image, global maxima, local maxima, moments etc. Statistical features can tolerate slight variations in signature style and distortion. And model training using deep learning techniques. Not been sufﬁciently addressed in the literature is the usage of one-class classiﬁcation models. One-class classiﬁcation systems that work well with low number of samples per user in deep learning.

T. Islam et.al [3] has introduced some simple objective measures of hesitation and investigating their reliability and efficacy in automating the analysis of handwritten signatures for forensic and biometric applications. Objective measure hesitation which is defined as number of times pen velocity falls beneath a threshold during single stroke and Feature extraction where some features such as distance travelled by pen ,number of pen lifts etc.. The issue with this paper is they didn’t focus on correlations between the hesitation feature and type of signature, its complexity.

S. Mushtaq et.al [4] has proposed to consider more reliable biometric feature, signature verification for the considering. The techniques are: verification performance measuring ratios such as The FRR is the ratio of the number of genuine test signatures rejected by the system to the total number of genuine test signatures submitted. The FAR is the ratio of the number of imposter accepted to the total number of forgeries submitted. Normalized signature area is the total number of signature pixels or foreground pixels in the signature image., Aspect ratio is defined as ratio of Width &Height. The accuracy attained so far from the existing systems is not very high and much more research on off-line signature verification is required. Future work may also include the fusion of different classifier for better verification results.

A. I. Al-Shoshan et.al [5] has proposed a dynamic signature verification system to meet the needs of the retail sector has been developed. The proposed system segments each signature based on its perceptually important points and then computes for each segment a number of features that are scale, rotation and displacement invariant. Dynamic Features and Fourier descriptor is a method of describing the shape of a closed, planar figure. The neural network used for pattern classification is done by structural analysis and Fourier descriptor.

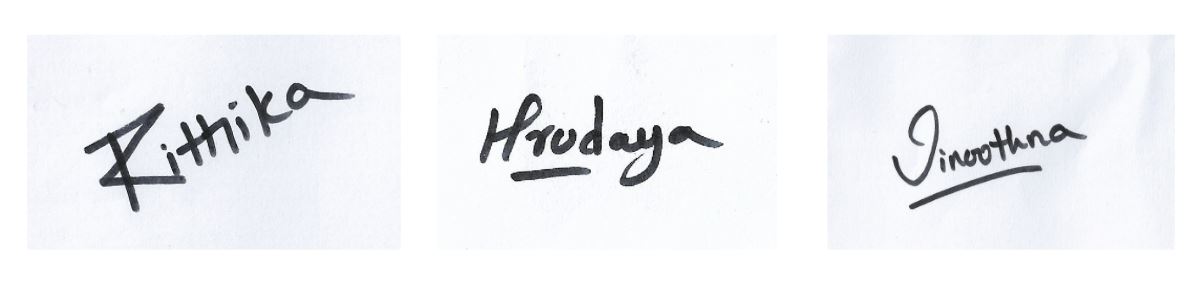
**CHAPTER 3**

**DATABASE**

Signature should be made on an unscrambled blank sheet of paper and shot with a camera which is focused on the image. The image should then be modified to a square resolution to be used.

**2.1 Database Preparation**

Signature database plays a significant role in the process of signature verification. In pattern classification, there are two phases - (i) training and (ii) testing. In training phase, the signature classifier is modeled. Sufficient numbers of signature samples are required to model the classifier. The larger database represents a population properly and it helps in producing more reliable results. Widely accepted and standard benchmark signature database is not yet available for offline signature verification. Therefore, we tried to develop our own signature databases for our experimentation. Figure 2.1 shows the signatures which are taken as database with five sets of each signature respectively.



**Figure 2.1** database images

**2.2 Scanner Resolution**

A scanner is specified by many factors such as - Scanner Type, Scanner Resolution, Maximum Resolution, Maximum Scan Area, Scanning Speed, Light Source, Color Bit Depth etc. Our concern was only the Scanner Resolution. Because other parameters do not affect the essential quality of the signature images required for our purpose. Scanner Resolution is a measurement of the resolving power of a scanner. Resolution of a scanner is the measurement of number of pixels that it can sample in the scanned image. It is expressed in dots per inch (dpi). 200 dpi means 200×200 or 40,000 dots per square inch. If resolution of a scanner is more, it implies that more numbers of dots or pixels are captured per inch of the image by the scanner. Image scanned with a higher resolution can be enlarged more. Thus scanner resolution indicates the enlargement capacity of the scanner. Resolution of scanned images is also an important factor that influences the process of signature verification. High resolution results more detailed images. But, they need more storage space and may contain noise. Thus computational cost is also higher. On the other hand, computational cost is lower with low resolutions. But required information may be lost in such images as their quality is affected by lower resolution. Therefore, resolution of the signature images should be suitably decided.

**2.3 Signature Preparation**

There must be a proper way of collecting the signature samples. Quality of the signature database largely relies on the signature collection protocol. To develop a high quality signature database for research purpose, their collection must be carefully controlled. It is seen that the genuine signatures possess high stability and less variation. But the forged signatures are highly inconsistent. Forged signatures are collected from lay forgers. It is not possible to convince (or even find) a professional forger to develop a database of forged signatures. It is difficult for a non-professional forger to replicate a signature as close as the genuine signature. Such forged signatures are not only deviated from their genuine counterpart, they contain large variations among themselves. These affect the performance of the signature verification system to overcome the inconsistency in the forged signatures, following approaches are suggested by the researches:

(i) The forger should be motivated by an award for his good consistent work and they must be encouraged to practice sufficiently.

(ii) Forgers with inconsistency should be denied while developing the database.

(iii) Stability of the signatures should be checked using statistical test. Signature samples passing a minimum predefined threshold limit should only be included in the database.

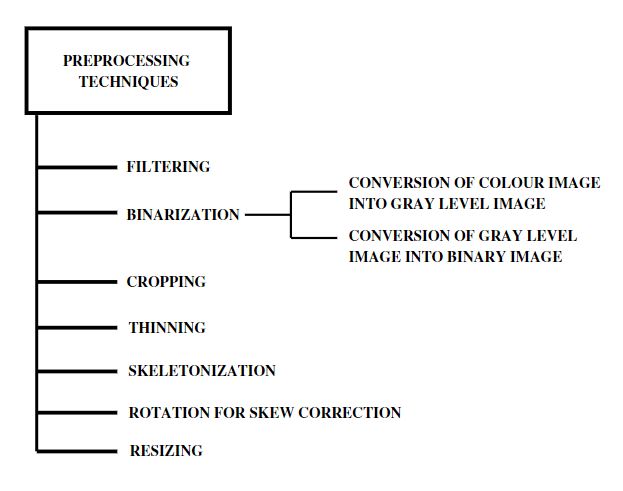
(iv) Some simulated signatures may be generated from the genuine signatures. These simulated signatures should be used as the forged signatures.

**CHAPTER 3**

**METHODOLOGIES**

**3.1 PREPROCESSING TECHNIQUES**

After capturing the signature samples, the next step is to enhance the images and make them ready for the subsequent processing. The images need to be preprocessed before giving them to the next process. Preprocessing is done using signal processing algorithms. Preprocessing greatly helps to improve the performance of feature extraction and classification. Depending on the type of signature pattern, signature image quality and classification techniques to be used, preprocessing operations are determined. It must be kept in mind that during preprocessing, information from the images should not be discarded. Loss of information in preprocessing will affect the overall accuracy of the signature verification system.

****

**Figure 3.1** preprocessing techniques

Figure 3.1 shows the steps involved in preprocessing techniques which are applied on the image from the database.

**3.1.1 Filtering**

A scanned signature image may contain noise. Noise in the image deteriorates the feature extraction and its successive processes. Hence, filtering of noise is an unavoidable preprocessing step in pattern recognition. It has been observed that the scanned images are usually affected by salt-peeper noise. A median filter effectively removes such type of noise preserving the edges of the images we applied a median filter of 3×3 window on our signature images. The median filter [13] is a non-linear spatial filter that uses a sub-image area or window. This window is usually of square shape and is of fixed size. This window slides over complete image pixel by pixel and replaces the center value in the window with the median of all the pixel values in the window.

The pixel value of the window in Fig. 3.2 in ascending order is 1, 4,5,5,6,6,7,10,67

|  |  |  |
| --- | --- | --- |
| 5 | 6 | 7 |
| 1 | 67 | 6 |
| 10 | 5 | 4 |

|  |  |  |
| --- | --- | --- |
| 5 | 6 | 7 |
| 1 | 5 | 6 |
| 10 | 5 | 4 |

**Figure 3.2** 3x3 window **Figure 3.3** window after noise removal

Figure 3.2 Shows the median is (the middle value of the string) 5. When the center value in the window (67) which is possibly a noise, is replaced with the median value (5), the following new window in Figure 3.3 is found, where the noise is removed.

**3.1.2 Binarization**

A color image comprises of three color plans Red (R), Green (G) and Blue (B). In a color image, every pixel value is defined by the combination of the values of these three plans. In a gray level image, there is no color information. The image is defined by the pixel values of a single plan (the intensity plan). However, in a gray level image, the pixel values will have a range which is specified by the number of bits of the image. Eg for an 8 bit image the pixel values will range from 0 to 255.

1. Conversion of Colour image into Gray level image
2. Conversion of Gray level image into Binary image

**3.1.2.1 Conversion of Color image into Gray level image**

There are several algorithms for converting a color image into a gray level image. The following four algorithms are found to be more common in used:

**i) Lightness method**

Here, the most prominent and least prominent colors are averaged. I= (max(R, G, B) +min(R, G, B))//2 **(3.1)**

I = Calculated gray level

R= Value of Red color plan

G= Value of Green color plan

B= Value of Blue color plan

**(ii) Average method**

In this method, the average value of R, G, B plan is considered as the gray level. I=R+G+B/3 **(3.2)**

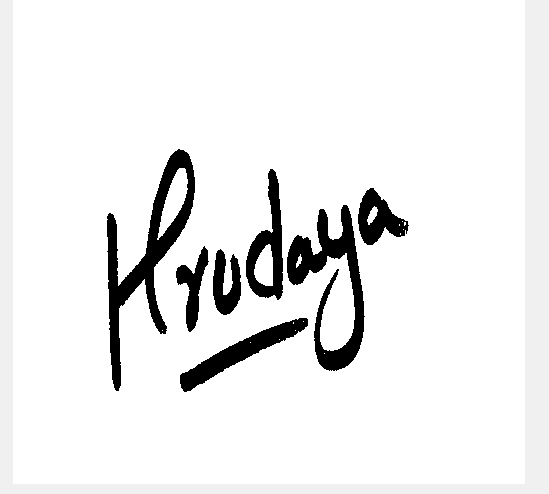
**(iii) Luminosity method**

Human perception is taken into account in this method. It is weighted average method. Human brain is more sensitive to green color than red and it is least sensitive to blue color. Accordingly in this method, different weights are given to these colors. I=0.21R+ 0.72G+ 0.07B **(3.3)**

**(iv) Standard NTSC (National Television System Committee) conversion Method**

Perception of this method is also similar to Luminosity method, but here the weights are different. This method is a standard method accepted by NTSC and is widely used. MATLAB® Image Processing Toolbox uses this method for converting a colour image into a gray level image. Figure 3.4 shows one of the preprocessing techniques where the color image is converted into gray level image.

I=0.2989R +0.587G+ 0. 114B **(3.4)**



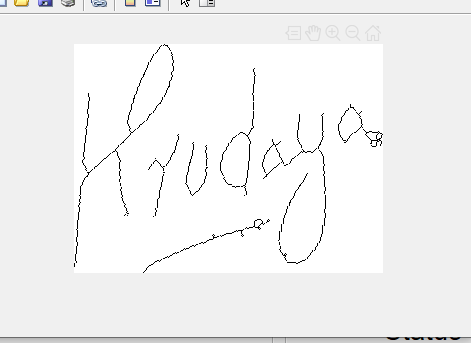
**Figure 3.4** conversion of color image into gray level image

**3.1.2.2 Conversion of Gray level image into Binary image**

A suitable threshold value (pixel value) is considered to convert a Gray level image into a binary image. If a pixel value in the gray level image is greater than the threshold value then the new pixel value assigned is 1 (one) else 0 (zero). Thus, the new image will have only two pixel values ‘1’ (which corresponds to white) and ‘0’ (which corresponds to black).

**3.1.3 Cropping**

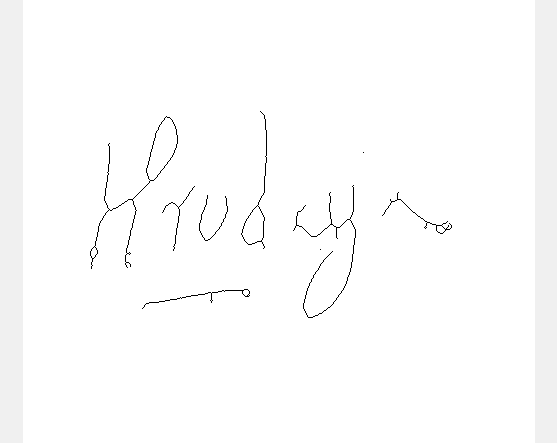
When scanned, signature image contains the signature and some white cultured non-signature regions. Those superfluous non-image portions are removed by cropping the image to the bounding rectangle of the signature part. Cropping is an essential preprocessing step for all types of classification techniques. Figure 3.5 shows one of the preprocessing techniques where the image is cropped to a bounding rectangle.

****

**Figure 3.5** cropped image

**3.1.4 Thinning**

In thinning, the signature image strokes are made one pixel thick. Thinning is mainly done to reduce the amount of data in the image. This helps to decrease the storage space requirement and also to reduce the computational complexities in successive stages. But during thinning, some information of the signature images such as stroke width may be lost. So, depending on the features to be extracted, thinning may or may not be required. Figure 3.6 shows one of the preprocessing techniques where the image is thinned for removing the foreground.

****

**Figure 3.6** thinned image

**3.1.5 Skeletonization**

Skeletonization is applied on binary images. It preserves the connectivity of the signature segments which were originally connected and removes selected foreground pixels from the image. After skeletonization, the signature image is converted into combination of some thin arcs and curves. One basic way of performing skeletonization is by using a morphological thinning process that successively removes pixels from the boundary. This process continues until no more thinning is possible.

**3.1.6 Rotation for Skew Correction**

Many a times, it is seen that during scanning of the signature images, the images are not properly oriented. This angular tilt in the signature image is called ‘skew’. Skew may result poor classification (depending on the classification technique used). Therefore there may be a need for skew correction of the signature images by rotating them. After skew correction, the final image is made parallel to the horizontal axis. In one method of skew correction, the bottom pixels of the signature image are used. A straight line is fitted through the bottom pixels by using polynomial curve fitting method. Skew angle of this line is measured and then skew correction is done by rotating the signature image by a negative skew angle. These are the steps involved in skew correction:

1. Move the signature to origin by using the co-ordinates of center of mass of the signature image.

2. Calculate the minimum Eigen Value of the matrix formed by using the new co-ordinates of the signature image.

3. Calculate the Skew angle using the Eigen vector

φ= [(M 1) / (M 2)]  **(3.5)**

Where,

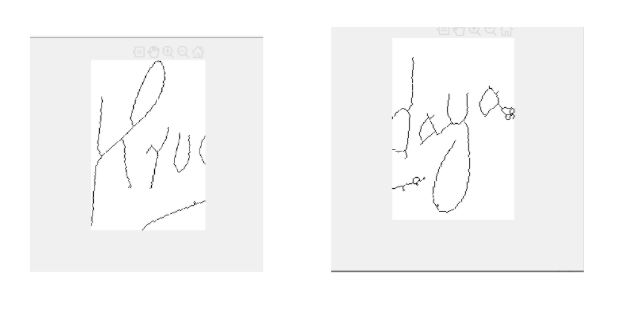
φ = Skew angle

M = the image matrix

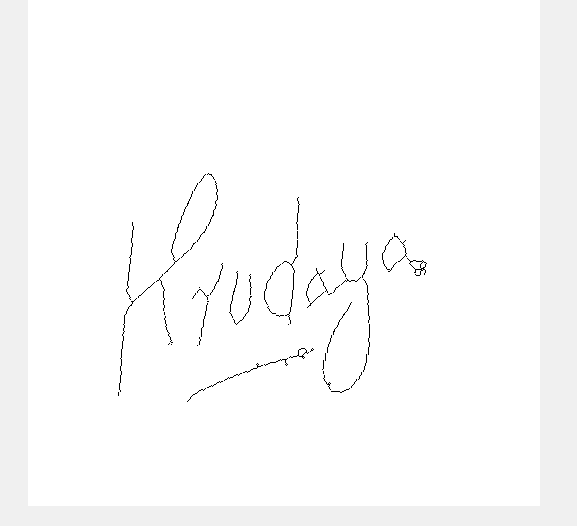
M (1) = the first element in the first row of the image matrix

M (2) = the first element in the second row of the image matrix

4. After skew angle is found, skew correction is performed by applying rotation transformation to every pixel in the signature image. Figure 3.7 and 3.8 shows the angular rotation of the images where every pixel of the image is rotated and in figure 3.9 the image is made parallel to the horizontal axis.



**Figure 3.7** central gravity (1)  **Figure 3.8** central gravity (2)

****

**Figure 3.9** image after skew correction

**3.1.7 Resizing**

Signature lengths are different for different signers. Even the lengths of the signatures of a single person are also not equal. But when a grid based signature verification approach is used, the signatures are projected on the grid of same size. Hence, all the signatures must be of same size. Therefore in that case, resizing of signature becomes important but, resizing is not a compulsory preprocessing step for all signature verification approaches. The most basic method of image resizing is a kind of geometric transformation. In this method, there are two basic operations: (i) spatial transformation and (ii) gray level interpolation.

In spatial transformation, some pixels or points (‘tie-points’) are selected whose positions in the original image and the resized image are precisely known. From their locations in the two images, a spatial transformation equation is formulated. This equation is used as a mapping equation to find out the positions of all the pixels in the new resized image.

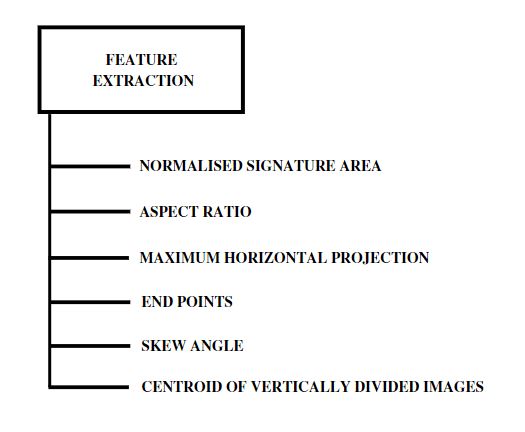
Gray level interpolation is used to assign gray levels to the new pixels in the resized image. It uses a nearest neighbor approach. In this method, gray level is assigned according to the pixel which is the nearest to the mapped pixel. Figure 3.10 shows one of the preprocessing techniques where the image is resized so that signatures projected are of same size.



**Figure 3.10** resized image

**3.2FEATURE EXTRACTION**

In image processing the main focal point of feature extraction is to obtain the most appropriate information from the input image. In the feature extraction, not only the important features are extracted, but also the surplus data is being eliminated. Transforming the image into set of features is called feature extraction. When the input image is too large to be processed then feature extraction banishes the undesirable features. Feature extraction can be used in many applications such as object recognition, data compression, pattern recognition etc. [14]



**Figure 4.1** feature extraction techniques

Figure 4.1 shows the steps involved in feature extraction which are applied on the image from the database after preprocessing it.

**3.2.1 Normalized Signature Area**

The purpose of using normalized signature area is to perceive the ratio of number of black pixels with the total number of pixels in the input image. The main motive for finding out only the black pixels is, the main part of the image in processing is signature in which the track of the image is a black pixel. Finding out the number of black pixels would make feature matching implementation easier.

To find out the number following methods can be implemented such as, transmuting the image in to its negative and finding out the number of pixels by the ratio between black and total number of pixels in the image. And the other is, the pixel is confirmed to be black if comparison with other intensity pixel is false .Find out the ratio between the obtained black pixels and total number of pixels gives the area of black pixel. This is one of the feature extraction processes to obtain the ratio of black pixels in the image.

Normalized Signature Area =number of pixels which belong to the signature /total number of pixels in the signature image

**3.2.2 Aspect Ratio**

Aspect ratio gives the relation between the image’s width and height. Aspect ratio may vary in line with the application or in accordance with the processing method. There are many standard aspect ratios. There are two types of aspect ratios namely, PAR (Pixel Aspect Ratio) and DAR (Display Aspect Ratio) [8]. In the project, the image stored in the database is a square image. Display Aspect Ratio of a square image is 1:1. Aspect ratio can be helpful in cropping. The amount of cropping done on an image can be decided by having a specific aspect ratio according to the particular application that the image is being used. The aspect ratio can be calculated with two parameters height and width; these parameters can be obtained by size function in the mat lab which gives the size of the image.

****

**Figure 4.2** aspect ratio

Figure 4.2 shows the aspect ratio which describes the width and height of the user’s signature.

**3.2.3 Maximum Horizontal Projection**

Maximum horizontal projection is to obtain the number of black pixels per a horizontal row of an image. The number of black pixels can be obtained by a piece of code comparing the reference pixel i.e., pixel of intensity value zero with all pixel intensities in the image and updating the count value. The other method to find black pixel is by using the mat lab function nnz. The nnz function [9] gives the number of zeroes in the matrix. By passing the input image as a parameter in the nnz function the number of pixels having intensity value of zero i.e., black pixels are obtained.

P h [y] =∑ black pixel(x, y) (from x=1 to n)  **(4.1)**

Pv [x] =∑ black pixel(x, y)  **(4.2)**

P h [y] = horizontal projection

Pv [x] = vertical projection

Where, m = width of the signature image

n = height of the signature image



**Figure 4.3** maximum horizontal projection

Figure 4.2 shows the number of black pixels per a horizontal row of the image from the database.

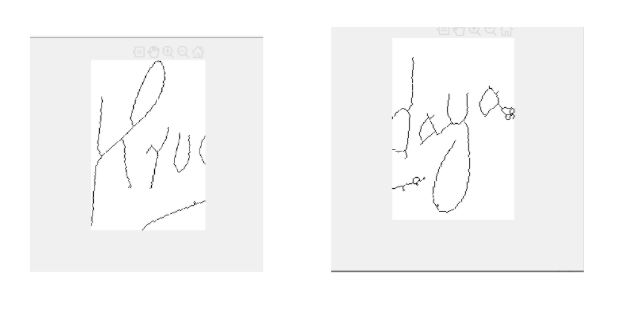
**Maximum horizontal projection (Maximum horizontal histogram):** In a horizontal projection, the row with maximum value gives the maximum horizontal histogram. **Maximum vertical projection (Maximum vertical histogram):** The highest value of the projection histogram in the vertical histogram is the maximum vertical projection.

**Centre of Gravity or Centroid:** In a binary signature image with black signature pixels, Centre of Gravity (CG) or Centroid is the average coordinate point of all black pixels.

**3.2.4 Centroid of Vertically Divided Images**

In a binary image or a black and white image with black signature pixels centroid is the average of all coordinate points of the pixel intensity with zero value .

Figure 4.4 the signature image is vertically divided into two halves. These halves have two centers of gravities of each respectively.



**Figure 4.4** centroid of vertically divided images

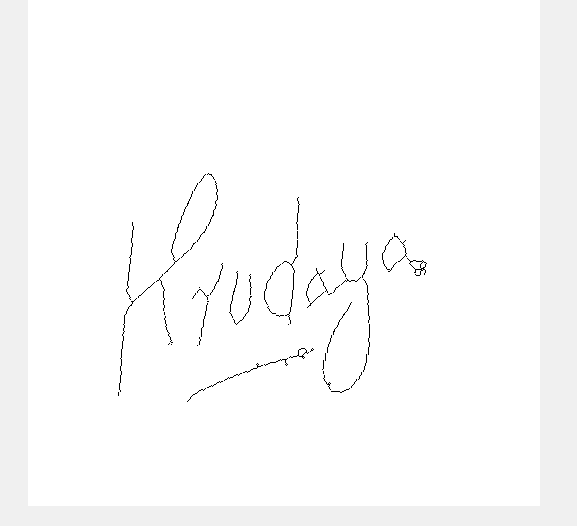
Figure 4.4 the signature image is vertically divided into two halves. These halves have two centers of gravities of each respectively.

**3.2.5 Skew Angle**

Skew angle is the angle between the normal plane of an image and the horizontal plane of the coordinate system. Skew angle aids to what amount the image should be rotated or shifted such that the image is properly aligned with the plane of coordinate system.

To align an image appropriate to the coordinate system, there three methods namely translation, rotation and scaling. Skew angle will categorize in rotation. There are two parameters to be known skew correction. The skewness of the image can be corrected by rotating the image at such an angle.

Many methods have been developed for the skew correction of images. Basically they are using projection profile using Hough transform technique Fourier method by nearest-neighbor clustering and correlation. In the projection profile method, a series of projections are obtained at a number of angles close to the expected orientation, and the variation is calculated for each of the projection. The profile that gives the maximum variation corresponds to the projection with the best alignment to the text lines. This projection angle is the skew angle.

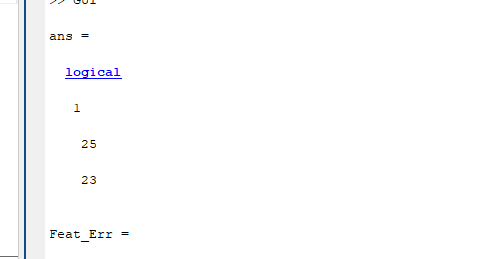


**Figure 4.5** skew angle

Figure 4.5 the signature image is rotated or shifted such that it is properly aligned with the plane of coordinate system.

**3.2.6 End Points**

The boundary pixels of an image are generally called as end points of an image. End points of an image can be detected by the neighborhood of the particular pixel. There are three types, namely 4-neighbourhood, 8-neighbourhood and diagonal neighborhood of a pixel. Edges can also be found out using neighborhood. If an edge does not have right side elements it is said to be a corner edge. Correspondingly, the end points can be distinguished by neighborhood. According to the coordinate location and neighborhood condition of a particular pixel in an image, the end points could be detected.



**Figure 4.6** end points

Figure 4.6 the signature image’s boundary pixels are shown.

**3.2.7 Number of cross points**

In a skeletonized signature image, Cross point is a signature pixel or point that has more than two 8-neighbours. (When pixels are considered on a grid, a pixel is surrounded by 8-neighbouring pixels; they are called its 8-neighbours)

**3.2.8 Choice of Features**

While considering the choice of features in an offline signature verification system, following two aspects need to be considered:

1. Types of features

2. Total number of features

**3.2.1.1 Types of features:**

Ofﬂine signature veriﬁcation [2] has been studied from many perspectives, yielding multiple alternatives for feature extraction. Broadly speaking, the feature extraction techniques can be classiﬁed as Static or Pseudo-dynamic, where pseudo dynamic features attempt to recover dynamic information from the signature execution process (such as speed, pressure, etc.). Another broad categorization of the feature extraction methods is between Global and Local features. Global features describe the signature images as a whole - for example, features such as height, width of the signature, or in general feature extractors that are applied to the entire signature image. In contrast, local features describe parts of the images, either by segmenting the image (e.g. according to connected components) or most commonly by the dividing the image in a grid (of Cartesian or polar coordinates), and applying feature extractors in each part of the image.

**Geometric Features:** Geometric features measure the overall shape of a signature. This includes basic descriptors, such as the signature height, width, caliber (height-to-width ration) and area. More complex descriptors include the count of endpoints and closed loops. Besides using global descriptors, several authors also generate local geometric features by dividing the signature in a grid and calculating features from each cell. For example, using the pixel density within grids

**Graph metric features [2]:** Forensic document examiners use the concepts of graphology and graphometry to examine handwriting for several purposes, including detecting authenticity and forgery. A subset of graph metric features that could be described algorithmically, and proposed a set of feature descriptors. They considered the following static features: Caliber the ratio of Height / Width of the image; Proportion, referring to the symmetry of the signature, Alignment to baseline describing the angular displacement to a horizontal baseline, and Spacing - describing empty spaces between strokes.

**Directional features:** Directional features seek to describe the image in terms of the direction of the strokes in the signature. This descriptor represents local shapes in an image by a histogram of edge orientations, also in multiple scales.

**Mathematical transformations:** [2]Researchers have used a variety of mathematical transformations as feature a grid is overlaid on top of the signature image, containing horizontal, vertical and diagonal bars, each bar containing a ﬁxed number of bins. Each pixel of the signature image is then projected to its closest bar in each direction, activating the respective bin. The count of active bins in the projections is then used as a descriptor of the signature. To achieve promising results on writer-independent and writer-dependent classiﬁcation, respectively.

**Texture features:** Texture features, in particular variants of Local Binary Patterns (LBP), have been used in many experiments in recent years. The LBP operator describes the local patterns in the image, and the histogram of these patterns is used as a feature descriptor. And demonstrated to be among the best hand-crafted feature extractors for this task. Another important texture descriptor is GLCM (Gray Level Co-occurrence Matrix).

**Interest point matching:** Interest point matching methods, such as SIFT (Scale-Invariant Feature Transform) and SURF (Speeded up Robust Features) have been largely used for computer vision tasks. After extracting interest points from both images, they generated a set of 12 features, using information such as the number of SIFT matches between the two images, and processing time During classiﬁcation, only the stable interest points are used for matching. The number of key points in the query image, and the number of matched key points were used to classify the signature as genuine or forgery.

**Pseudo-dynamic features:** A set of pseudo-dynamic features, based on graph metric studies: Distribution of pixels, Progression - that measures the tension in the strokes, provides information about the speed, continuity and uniformity, Slant and Form - measuring the concavities in the signature. More recently, this was accomplished by ﬁtting Benzier curves to the signature outline (more specially, to the largest segment of the signature), and using the parameters of the curves as features.

**3.2.2 Total number of features:**

In pattern recognition, feature selection is a process of selecting a subset of the most important and relevant features from the complete set of extracted features so that the total number of features is reduced but their class discriminatory information is still intact .All the extracted features carry information of the pattern from which they were extracted. But all of them don’t carry information about the pattern equally. In a feature set, some of the features are irrelevant or redundant. If the classifier is fed with the irrelevant features, there three problems may be induced by the irrelevant features:

(i) Due to more number of features, computational cost increases.

(ii) Presence of the irrelevant features may cause misclassification and thus the classification efficiency is decreased.

**CHAPTER 5**

**FEATURE MATCHING**

Feature matching is an important task in many computer vision applications, such as structure-from-motion, image retrieval, object detection, and more.

A feature is a piece of information which is relevant for solving the computational task related to a certain application. Features may be specific structures in the image such as points, edges or objects. Features may also be the result of a general neighborhood operation or feature detection applied to the image. The features can be classified into two main categories:

* The features that are in specific locations of the images, such as mountain peaks, building corners, doorways, or interestingly shaped patches of snow. These kinds of localized features are often called **key point features**(or even corners) and are often described by the appearance of patches of pixels surrounding the point location.
* The features that can be matched based on their orientation and local appearance (edge profiles) are called **edges** and they can also be good indicators of object boundaries and occlusion events in the image sequence.

**5.1 Components of Feature Detection and Matching**

* **Detection:** Identify the **Interest Point.**
* **Description:** The local appearance around each feature point is described in some way that is (ideally) invariant under changes in illumination, translation, scale, and in-plane rotation. We typically end up with a descriptor vector for each feature point.
* **Matching:** Descriptors are compared across the images, to identify similar features. For two images we may get a set of pairs (***Xi, Yi***) ↔ (***Xi`, Yi`***), where (***Xi, Yi***) is a feature in one image and (***Xi`, Yi`***) its matching feature in the other image.

**5.1.1 Interest Point**

Interest point or Feature Point is the point which is expressive in texture. Interest point is the point at which the direction of the boundary of the object changes abruptly or intersection point between two or more edge segments.

**Properties of Interest Point:**

* It has a well-defined *position* in image space or well localized.
* It is *stable* under local and global perturbations in the image domain as illumination/brightness variations, such that the interest points can be reliably computed with a high degree of *repeatability*.
* Should provide efficient detection.

**Possible Approaches**

* Based on the brightness of an image (Usually by image derivative).
* Based on Boundary extraction (Usually by Edge detection and Curvature analysis).

**5.1.2 Feature Descriptor**

A feature descriptor [5] is an algorithm which takes an image and outputs feature descriptors/feature vectors. Feature descriptors encode interesting information into a series of numbers and act as a sort of numerical “fingerprint” that can be used to differentiate one feature from another.

Ideally, this information would be invariant under image transformation, so we can find the feature again even if the image is transformed in some way. After detecting interest point we go on to compute a descriptor for every one of them. Descriptors can be categorized into two classes:

* **Local Descriptor:** It is a compact representation of a point’s local neighbourhood. Local descriptors try to resemble shape and appearance only in a local neighbourhood around a point and thus are very suitable for representing it in terms of matching.
* **Global Descriptor**: A global descriptor describes the whole image. They are generally not very robust as a change in part of the image may cause it to fail as it will affect the resulting descriptor.

**5.1.3 Feature Matching**

Features matching or generally image matching, a part of many computer vision applications such as image registration, camera calibration and object recognition, is the task of establishing correspondences between two images of the same scene/object. A common approach to image matching consists of detecting a set of interest points each associated with image descriptors from image data. Once the features and their descriptors have been extracted from two or more images, the next step is to establish some preliminary feature matches between these images.

Generally, the performance of matching methods based on interest points depends on both the properties of the underlying interest points and the choice of associated image descriptors. Thus, detectors and descriptors appropriate for images contents shall be used in applications. For instance, if an image contains bacteria cells, the blob detector should be used rather than the corner detector. But, if the image is an aerial view of a city, the corner detector is suitable to find man-made structures. Furthermore, selecting a detector and a descriptor that addresses the image degradation is very important.

**5.2 Threshold Calculation and Matching**

Local maxima and minima of difference-of-Gaussian D are computed by comparing the sample point with its eight neighbors in the current image as well as the nine neighbors in the scales above and below (total of 26 neighbors). If the pixel represents a local maximum or minimum, it is selected as a candidate key point. The adaptive threshold method is applied after key point extreme calculation. When all the extrema are calculated, corresponding Gaussian values to the key points are compared with the threshold calculated for brighter and darker region. Then two criteria are used for the detection of unreliable key points. The first criteria evaluate the value of IDI at each candidate key point. If the value is below certain threshold, which means that the structure has low contrast, the key point is removed. The second criteria evaluate the ratio of principal curvatures of each candidate key point to search for poorly defamed peaks in the Difference-of- Gaussian function.

Firstly, for selection of threshold for darker and brighter regions, a boundary (middle point) is selected between darker and brighter pixels. For example, in a Gaussians of an image we have pixel values from 0 to 1. We select middle pixel value of 0.5 as a boundary. The brighter threshold is selected by taking the mean Gaussian value of brighter pixels. Similarly darker threshold is selected by calculating the mean Gaussian value for darker pixels. Darker pixel values vary from 0 to 0.5 and brighter pixel values vary from 0.5 to 1. Key points in the darker region of image are compared with darker threshold. If key point is smaller than darker threshold then it is selected otherwise it is rejected. Similarly, key points in brighter region of image are compared with brighter threshold. Likewise, the key points are selected only if they are greater than this brighter threshold.

In the matching phase, the key points with darker region of test image are matched separately with key points of darker region of training image. Similarly, brighter region of test image is matched with brighter region of training image to avoid false matches due to illumination changes. If user requires changing the brighter and darker region for including some specific key points, then it can be performed by changing the boundary (midpoint) pixel value of darker and brighter region. Experiment results show that our approach works well in many cases. Most key points are present either in darker or in brighter region and only few key points are along the boundary of brighter and darker regions in image. So selecting two thresholds adaptively for brighter and darker regions collects the useful key points in both regions and removes the effect of mismatches. For example, if a key point is in brighter region of the training image and is in the darker region of the test image, then the by SIFT method it can make them a correct match. But in our method, two adaptive thresholds remove such mismatch errors and key points are matched with improved accuracy in specific regions.

**CHAPTER 6**

**DECISION FOR AUTHENTICATION**

The Decision for Authentication procedure passes through the feature space to find out the best feature subset. Various algorithms are used for the Decision for Authentication. Some popular search algorithms are: greedy hill climbing, Exhaustive, Best first, Simulated annealing, Genetic algorithm, Greedy forward selection, Greedy backward elimination, Particle swarm optimization, Targeted projection pursuit, Scatter Search, Variable Neighborhood Search etc. Following are some of the search methods available in WEKA data mining software which have been used in our research work.

**6.1 Greedy Hill Climbing Search Algorithm**

Only the local changes to a feature subset are considered in greedy hill climbing search algorithm. This algorithm [15] evaluates the resulted local changes in the present feature set by adding or removing one feature. The change which improves the score is chosen. A change made once for a feature subset is not considered again.

**6.2 Best First Search Algorithm**

This algorithm is similar to that of greedy hill climbing algorithm. But it tracks back to a previous subset if the current subset has not been found to be promising. Best first performs greedy hill climbing with back tracking. The system backtracks after it encounters a pre-specified number of non-improving nodes. The search process may start from an empty set or from a full set of features or it may start from an intermediate point. Search for the best features are made in either direction by considering all possible single feature additions and deletions.

**6.3 Exhaustive Search Algorithm**

This algorithm [15] searches for every possible subset that scores a minimum specified threshold. Sometimes, under some special requirements this threshold may be set to zero. At a defined stopping point, the highest scoring subset is selected as the best subset. The stopping criterion may be minimum threshold score, a maximum allowed run time etc.

**6.4 Ranker Search**

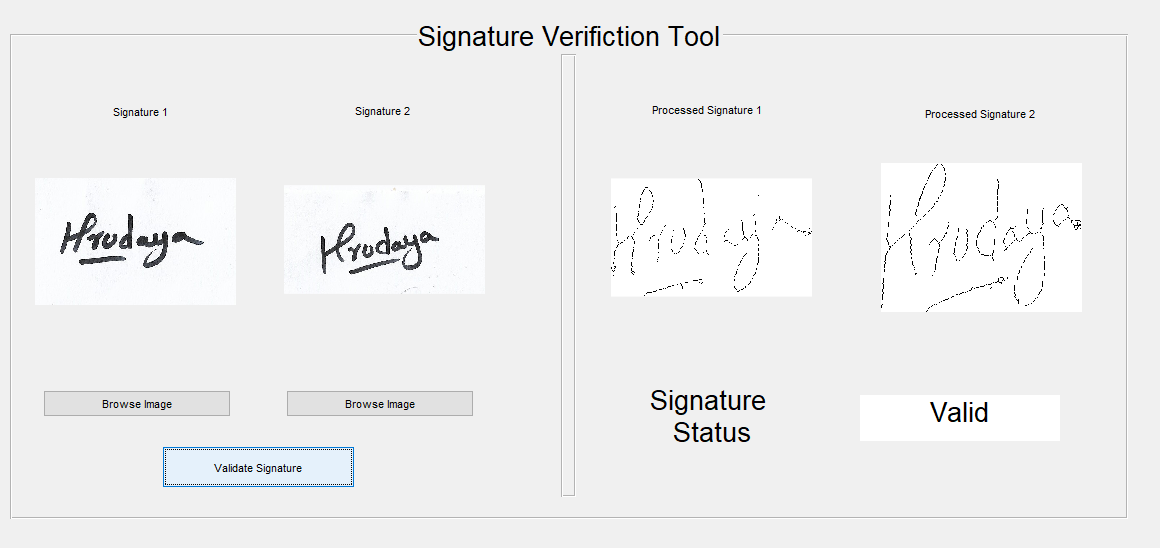
This search algorithm [15] uses a single feature evaluator and sorts the features and the potential feature subsets are ranked. At first, using the single feature evaluator, features are sorted and then using the subset evaluator, subsets of increasing size are evaluated. In this method, computational complexity is less and this method is very fast.

**6.5 Ranker**

This is not essentially a search method. This is a scheme used to rank the individual features. It is used with a feature subset evaluator. Every feature is evaluated individually. Ranks are given to the features based on their evaluation score. A threshold can be set to remove some of the lower-ranked features. Also one can specify the number of features to be selected after ranking. There is another option to keep hold of some features not considering their rank.

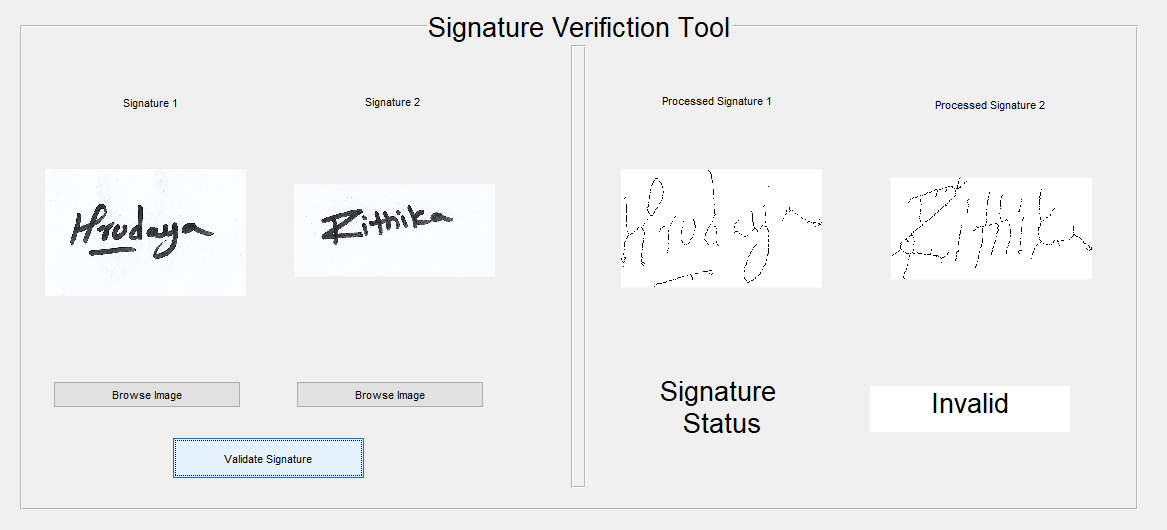
**6.6 Stop Point of Decision for Authentication**

A search algorithm goes on adding (if it is using forward selection) or removing (if it is using backward selection) features by continuously checking the score based on classification error. This process of removal or addition of features is stopped when it is not improving the score of the feature subset. Another way is to revise the feature subset continuously till its score doesn’t go down or to keep on generating the feature subsets and then the best feature subset is selected. Out of the different stopping criteria used to stop the search procedure, following two criteria are more commonly found to be in use: search stops when the minimum node size (minimum number of features) is reached and when the maximum tree depth is attained (i.e. maximum number of nodes are reached). The search procedure is terminated when the condition defined in the stopping criteria is satisfied. The stop criteria are guided by the search algorithm used.



**Figure 6.1** features of the signatures are matched

Figure 6.1 since the features of the signatures compared are matched, the signature status is shown as valid.



**Figure 6.2** features of the signatures are unmatched

Figure 6.2 since the features of the signatures compared are unmatched, the signature status is shown as invalid.

**CHAPTER 7**

**CONCLUSIONS**

A novel signature verification method was introduced, which is able to preserve and take usage of semantic information during signature comparison. Experimental results are promising, even now, when the algorithm uses only a fragment of this semantic information. Further benefits of this approach are, that each decision made by the algorithm, can be directly interpreted by humans, which facilitates the improvement and debugging of the whole system.

Therefore, an effective signature recognition and verification system has been built in order to draw a comparison between valid and invalid signatures. This system can be used in various applications like land documents, cheques, etc. to overcome forgery which takes place in huge number in day to day life.

The future development will mainly focus on the extraction and comparison of further features. Splines could be fit to the matched points and an appropriate metric could be found to measure the differences of the curves. Our experiences show that tremor is a very typical feature for forged signatures. This tremor could be measured using a Fourier transformation for calculating the distance from the smoothed curves. The endpoints could be also extended by a new metric, describing the runoff of the corresponding stroke. These runoffs could be characterized by the changes in the intensity. Threshold values in our system could be refined by loading the results from the different feature evaluations into an expert system.

**REFERENCES**

1. K. Krishnaditya, K. Varun, and K. Mohit, “Handwritten Signature Recognition: A Convolutional Neural Network Approach”, International Conference on Advanced Computation and Telecommunication, Bhopal, India, 2018, pp. 1-5, doi: 10.1109/ICACAT.2018.8933575
2. L. G. Hafemann, R. Sabourin and L. S. Oliveira, "Offline handwritten signature verification — Literature review," 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QC, 2017, pp. 1-8, doi: 10.1109/IPTA.2017.8310112
3. T. Islam and M. Fairhurst, "Investigating an objective measure of writer hesitation for forensic analysis of the handwritten signature," 2016 4th International Conference on Biometrics and Forensics (IWBF), Limassol, 2016, pp. 1-6, doi: 10.1109/IWBF.2016.7449686.
4. S. Mushtaq and A. H. Mir, "Signature verification: A study," 2013 4th International Conference on Computer and Communication Technology (ICCCT), Allahabad, 2013, pp. 258-263, doi: 10.1109/ICCCT.2013.6749637.
5. A. I. Al-Shoshan, "Handwritten Signature Verification Using Image Invariants and Dynamic Features," International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06), Sydney, Qld., 2006, pp. 173-176, doi: 10.1109/CGIV.2006.52
6. Y. M. Al-Omari, S. N. H. S. Abdullah and K. Omar, "State-of-the-art in offline signature verification system," 2011 International Conference on Pattern Analysis and Intelligence Robotics, Putrajaya, 2011, pp. 59-64, doi: 10.1109/ICPAIR.2011.5976912.
7. I. Abroug and N. E. Ben Amara, "Off-line signature verification systems: Recent advances," International Image Processing, Applications and Systems Conference, Sfax, 2014, pp. 1-6, doi: 10.1109/IPAS.2014.7043294.
8. U. A. Jain and N. N. Patil, "A comparative study of various methods for offline signature verification," 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, 2014, pp. 760-764, doi: 10.1109/ICICICT.2014.6781376.
9. O. o-khalifa, M. K. Alam and A. H. Abdalla, "An evaluation on offline signature verification using artificial neural network approach," 2013 INTERNATIONAL CONFERENCE ON COMPUTING, ELECTRICAL AND ELECTRONIC ENGINEERING (ICCEEE), Khartoum, 2013, pp. 368-371, doi: 10.1109/ICCEEE.2013.6633964.
10. B. Schafer and S. Viriri, "An off-line signature verification system," 2009 IEEE International Conference on Signal and Image Processing Applications, Kuala Lumpur, 2009, pp. 95-100, doi: 10.1109/ICSIPA.2009.5478727.
11. M. Hanmandlu, A. B. Sronothara and S. Vasikarla, "Deep Learning based Offline Signature Verification," 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York City, NY, USA, 2018, pp. 732-737, doi: 10.1109/UEMCON.2018.8796678.
12. D. R. Kisku, A. Rattani, P. Gupta and J. K. Sing, "Offline signature verification using geometric and orientation features with multiple experts fusion," 2011 3rd International Conference on Electronics Computer Technology, Kanyakumari, 2011, pp. 269-272, doi: 10.1109/ICECTECH.2011.5942000.
13. R. E. Twogood and F. G. Sommer, "Digital Image Processing," in IEEE Transactions on Nuclear Science, vol. 29, no. 3, pp. 1075-1086, June 1982, doi: 10.1109/TNS.1982.4336327.
14. S. Vanjimalar, D. Ramyachitra and P. Manikandan, "A Review on Feature Selection Techniques for Gene Expression Data," 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Madurai, India, 2018, pp. 1-4, doi: 10.1109/ICCIC.2018.8782294.
15. K. Naidu, A. Dhenge and K. Wankhade, "Feature Selection Algorithm for Improving the Performance of Classification: A Survey," 2014 Fourth International Conference on Communication Systems and Network Technologies, Bhopal, 2014, pp. 468-471, doi: 10.1109/CSNT.2014.99.