



Graph Based Offline Signature Recognition System by using Euclidian Distance

Ravikumar B. Chawhan

Department of computer science and engineering, Samskruti College of Engineering & Technology –Hyderabad

Abstract— Signature recognition is an important requirement of automatic document recognition system. Many approaches for signature recognition are found in literature. A novel approach for Graph based offline signature recognition system is presented in this paper, which is based on powerful Graph based features. The proposed system functions in three stages. Pre-processing stage; which consists of six steps: gray scale conversion, noise removal, normalization, binarization, resize, thinning to make signatures ready for feature extraction, Feature extraction stage; where totally 22 features are extracted which are used to distinguish the different signatures. Finally in classification stage; a simple Euclidean distance measure is used as decision tool. The average recognition accuracy obtained using this model ranges from 94% to 95% with the training set of 50 persons.

Keywords---Pre-processing; Binarization; Normalization; Euclidean Distance; Graph theory; Graph based features.

I. INTRODUCTION

“Biometrics” means “life measurement”. This term is generally couple with the use of unique and accurate physiological characteristics to identify a person. Some characteristics of biometrics are universal, invariance of properties, measurability, singularity, acceptance, reducibility, reliability. The advantage of biometrics is that, the information is unique for each individual and it can identify the individual in spite of variations in the time. Some of the examples of biometric systems are eyes-iris, retina, **face, fingerprint, finger geometry recognition**. Signature recognition is a behavioural biometric, It is characterized by a behavioural trait that is learnt and acquired over a period of time rather than physiological characteristics. Comparing to signature recognition system other biometric systems equipments are costly.

Signature recognition is special case of cursive handwriting recognition. Moreover, Performing cursive handwriting recognition or signature recognition is subjective rather than being objective. Signature recognition is designed to recognize subjects based on the traits of their unique signature. Signature has been a distinguishing feature for person identification. Even today, an increasing number of transactions, especially related to financial and business are being authorized via signatures. Hence, there is need to develop methods of automatic signature recognition system.

There are two types of handwritten signature recognition system one is online or dynamic signature recognition system another one is off-line or static signature recognition system. In online recognition, users write their signature in a digitizing tablet, which acquires the signature in real time. Dynamic signatures can be authenticated either at time of signing or post signing, and as triggers in workflow processes. Some systems also operate on smart-phones or tablets with a capacitive screen, where users can sign using a finger or an appropriate pen. Another one Off-line is complex due to the absence of stable dynamic characteristics. Difficulty also lies in the fact that it is hard to segment signature strokes due to highly stylish and unconventional writing styles. The objective of signature recognition is to discriminate between many persons, which are related to intra and interpersonal variability. Offline signature recognition is so different with the online recognition, because signature is often unreadable, and it seems it is just an image with some particular curves that represent the writing style of the person. Offline signature is just a special case of handwriting and often is just a symbol. So it is wisdom and necessary to just deal with a signature as a complete image with special distribution of pixels and representing a particular writing style and not as a collection of letters and words.

1.1 Graph Theory

Graph is the main concept of the graph theory. In our case, undirected weighted graphs will be used. An undirected mathematical graph G is an ordered pair (V, E) in which V represents a set of vertices (nodes) and E subset $V * V$ is a set of unordered pairs from V called set of edges of the graph G . In other words, mathematical graph is a set of vertices that are connected by links called edges. Every edge connects only two vertices, and every two vertices can be called adjacent only if they are connected with an edge.

1.2 Signature's graph creation

Signature graph is based on the number of the key points in the signature. Each stroke is a line and has two characteristic points. Therefore, each signature must have at least one stroke. Example of a stroke and its characteristic points. Number of point is an integer number and it will be very important in the process of creating a signature's graph. Characteristic points of each stroke will represent graph vertices and will be connected with and edge. But, before creating a graph it can be divided into segments according to signature's width.

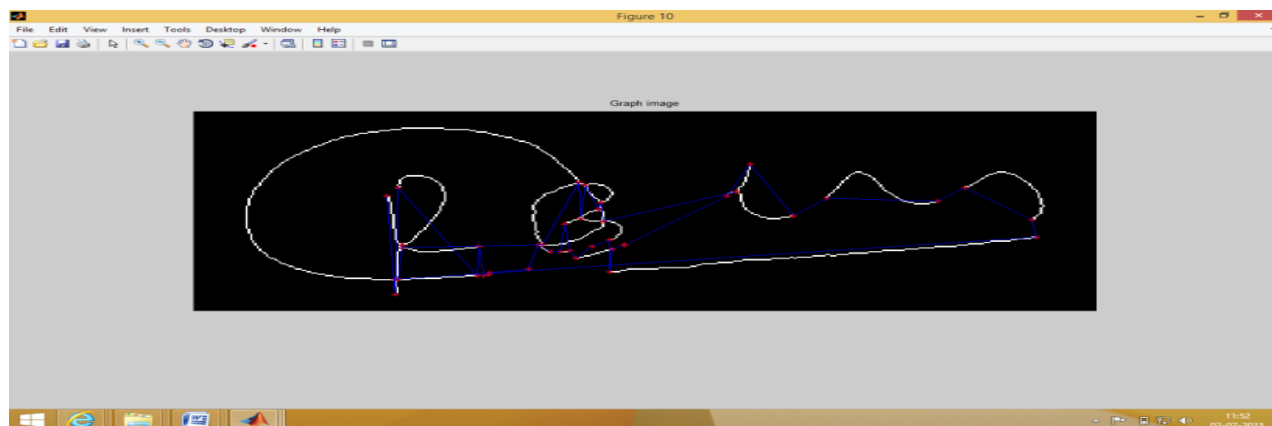


Figure: 1 Example of signature graph

1.3 Applications

Offline signatures are used for authentication, non-repudiation and integrity. It is also software distribution, financial transactions, digital document and in other cases where it is important to detect forgery or tampering. It is also used in e-tendering systems, wireless networks.

1.4 Motivation

Signature has been a distinguishing feature for person identification. Today an increasing number of transactions, especially related to financial and business are being authorized via signatures. The feature identification is very difficult in off-line signature recognition. In online signature where dynamic aspects of the signing action are captured directly as the handwriting trajectory. This motivated to build a signature recognition system which can solve the authentication problem up to some extent by using graph based features.

II. LITERATURE SURVEY

This literature survey [1]-[10] presents a brief survey of various offline approaches used to develop more robust and more constructive offline signature recognition approach. In this survey various classifiers methods to recognize the signature of the person are discussed. Some of them are support vector machine, neural networks (nn), simple distance classifier, euclidian distance classifier, hidden markove model (hmm) methods etc. Many online signature recognition/verification systems use graph based techniques since dynamic aspects of the signing action are captured directly as the handwriting trajectory. Graph based features in online systems have shown satisfactory results. This motivated to use the graph based techniques in offline signature recognition systems which can solve the authentication problem up to some extent. Euclidian distance is the simple and efficient classifier that can implement in offline signature recognition system. Graph based features are the strong features that can be extracted from a signature. That is why combination of these two is taken to develop a signature recognition system.

III. PROBLEM DEFINITION

The issues of signature recognition systems are variation in dimensions from one instance to another. There is a non repetitive nature of variation of the signature, because of age, illness. To overcome this problem we propose a graph based signature recognition. The problem definition defined after the related work “**Graph Based Signature Recognition System**”. The main aim of the work is to design and develop an offline handwritten signature recognition system that will differentiate between signatures using graph theory.

IV. PROPOSED METHODOLOGY

The methodology for signature recognition consists of some crucial steps which should be followed with specific rules. The simple block diagram shown below brings the brief idea about it.

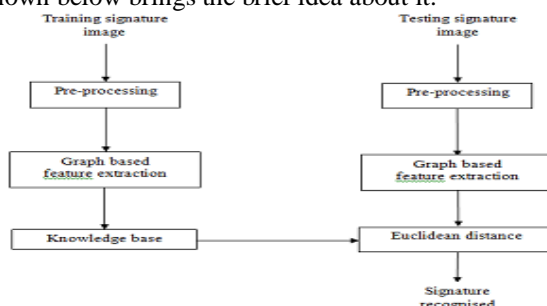


Figure: 2 Block Diagram of Proposed Work of Graph based signature recognition system

4.1 Image Acquisition

Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source. The signature database consists of a total of 500 handwritten signature images. Users write their signature on paper, digitize it through an optical scanner or a camera. These signatures were obtained from 50 persons each of them contributing 10 signatures.

4.2 Pre-Processing

Image Pre-Processing is a technique to enhance raw images by removing distortions and is the first part of the proposed system prior to computational processing. RGB-to-grayscale conversion, binarization, smoothing, resizing and thinning for pre-processing are applied in this stage. The following are the preprocessing methods under study.

- a) RGB to Gray-Scale Conversion
- b) Noise removal
- c) Normalization
- d) Binarization
- e) Resizing
- f) Thinning



Figure 4.1: (a) RGB Image (b) Gray-Scale Image

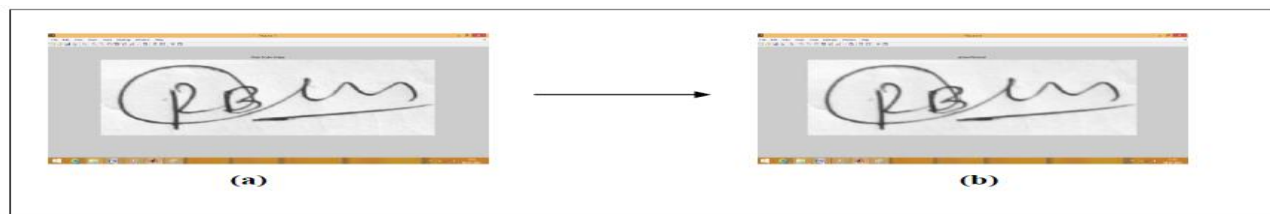


Figure 4.2: (a) Gray-Scale Image (b) Noise removal Image

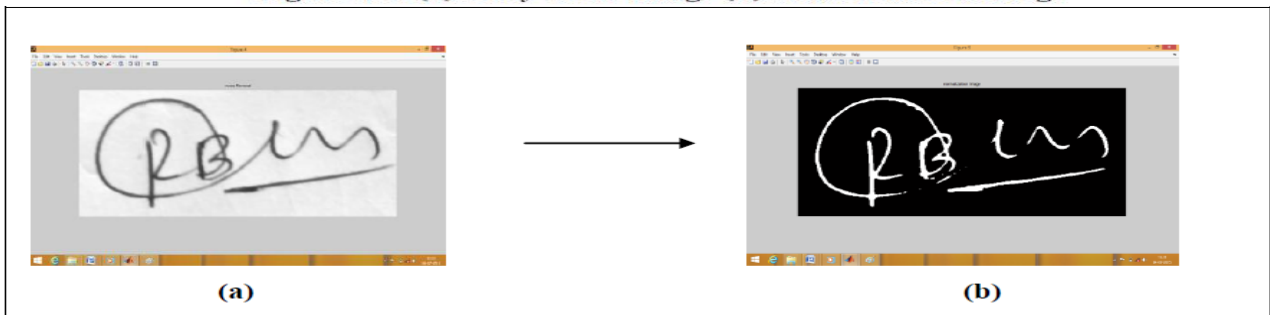


Figure 4.3(a) Noise removal Image (b) Normalization Image

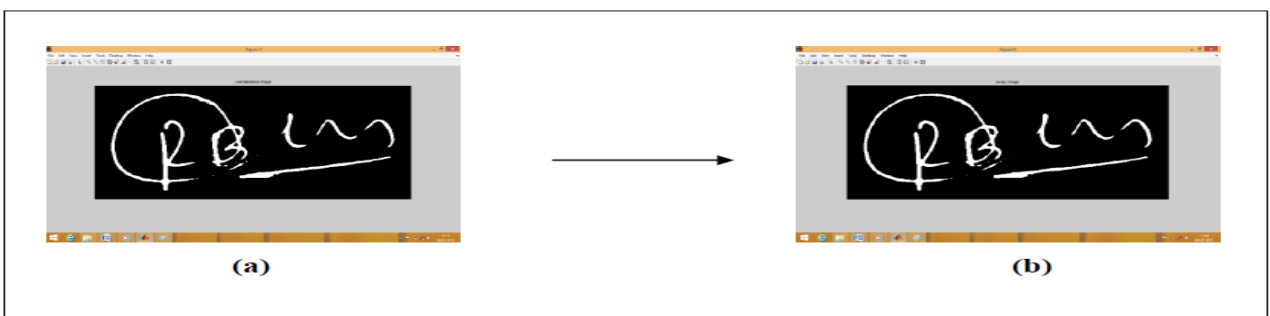


Figure 4.4: (a) Normalization Image (b) Binary Image

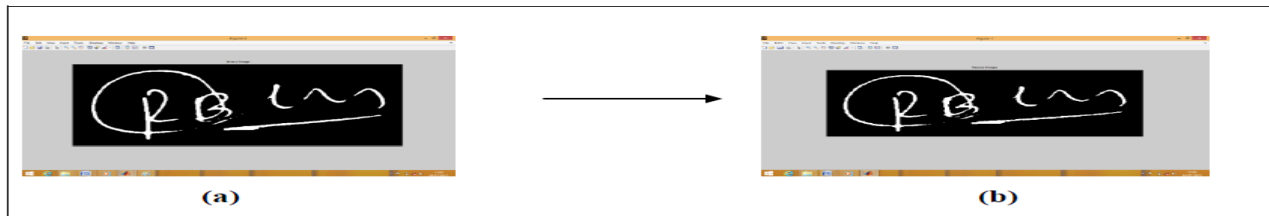


Figure 4.5: (a) Binary Image (b) Resize Image

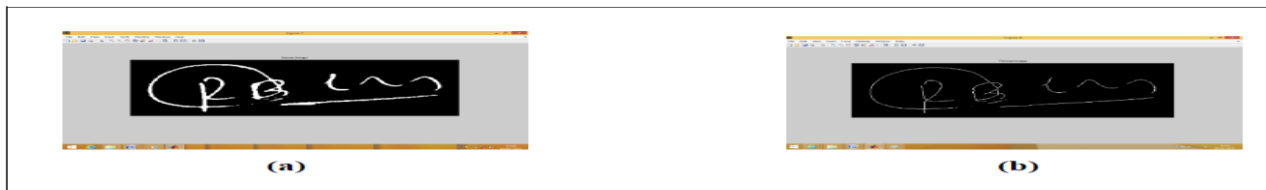


Figure 4.6 (a) Resize Image (b) Thinning Image

V. FEATURES EMPLOYED

The feature extraction method is most important step in any recognition system because the recognition accuracy totally depends on the feature extracted. The main objective of a feature extraction technique is to accurately retrieve the features. Graph based parameters such as nodes, edges, branch points and extreme points are extracted. The signature graph and graph image without signature shown in figure 3 and 4.

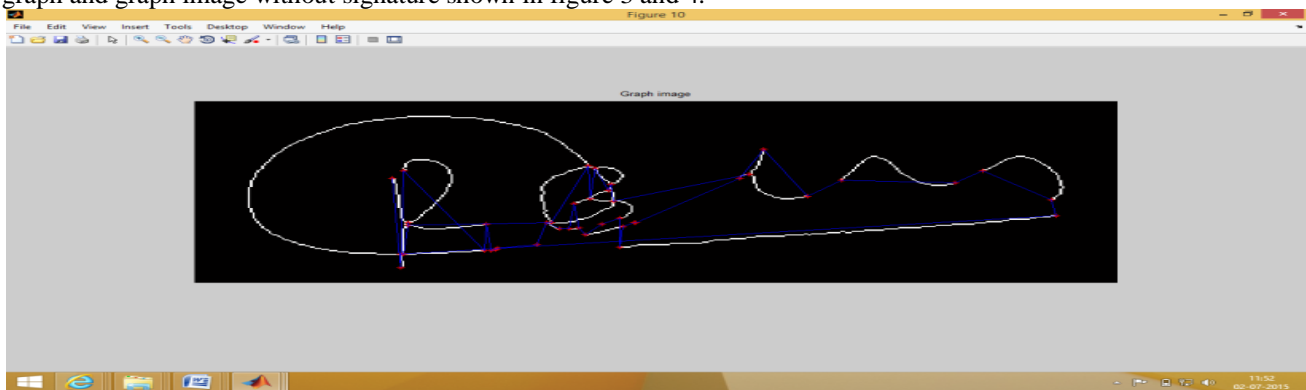


Figure: 3 Example of signature graph

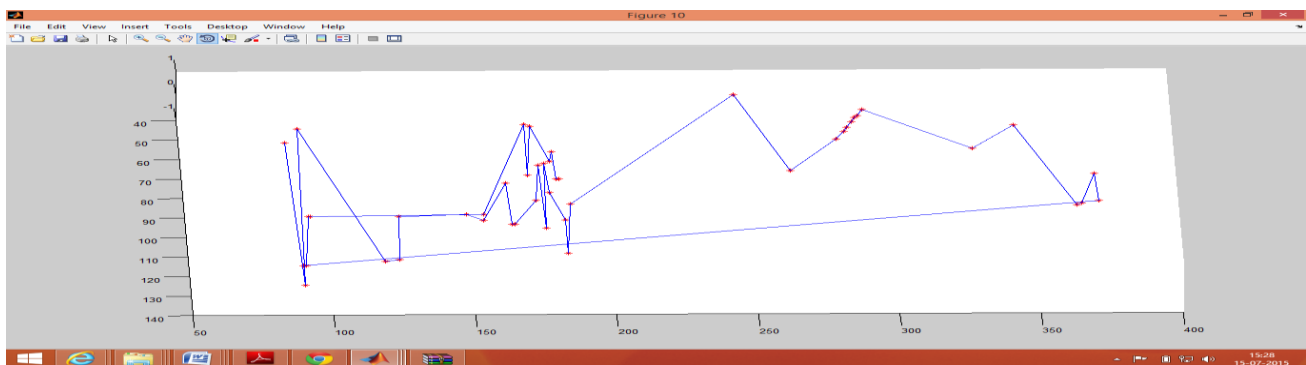


Figure: 4 Graph image without signature

5.1 Branch end points and Intersection points

End points are defined as those pixels, which have only one neighbour in its surroundings. The line connected with the end point is called end point line. The nature of the line could be an important feature. Some of end point line have same density pixel throughout, But some end point line has different pixel density (generally decreasing order). A branch point of an analytic function is a point in the complex plane whose complex argument can be mapped from a single point in the domain to multiple points in the range.

Intersection points for lines or polygon edges. The point or set of points where one line, surface or solid crosses another. mathematics the set that contains only those elements shared by two or more sets.

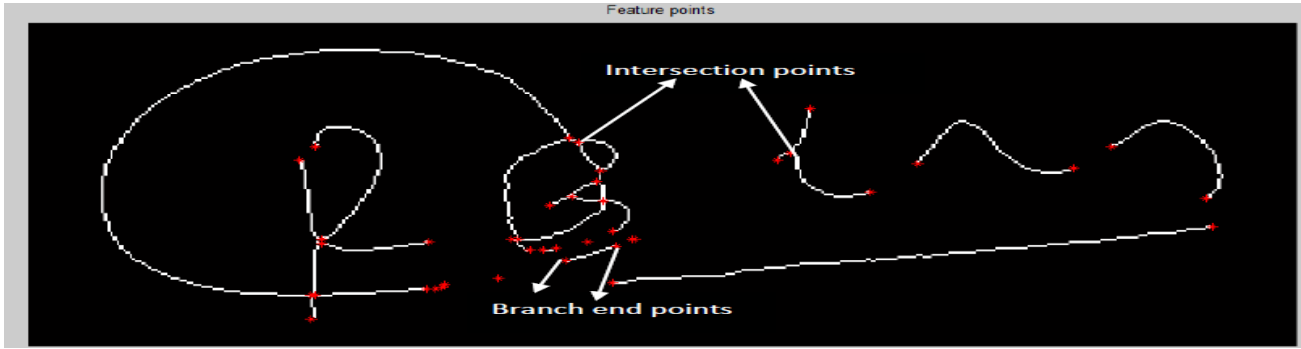


Figure: 5 Branch and intersection points

5.2 Extreme points, x, y co-ordinates, edge and node points

Extreme 8-by-2 matrix that specifies the extreme points in the region. Each row of the matrix contains the x- and y-coordinates of one of the points. The format of the vector is top-left, top-right, right-top, right-bottom, bottom-right, bottom-left, left-bottom, left-top. This property is supported only for 2-D input label matrices. The extreme points illustrates of two different regions. In the region on the left, each extreme point is distinct. In the region on the right, certain extreme points (e.g., top-left and left-top) are identical.

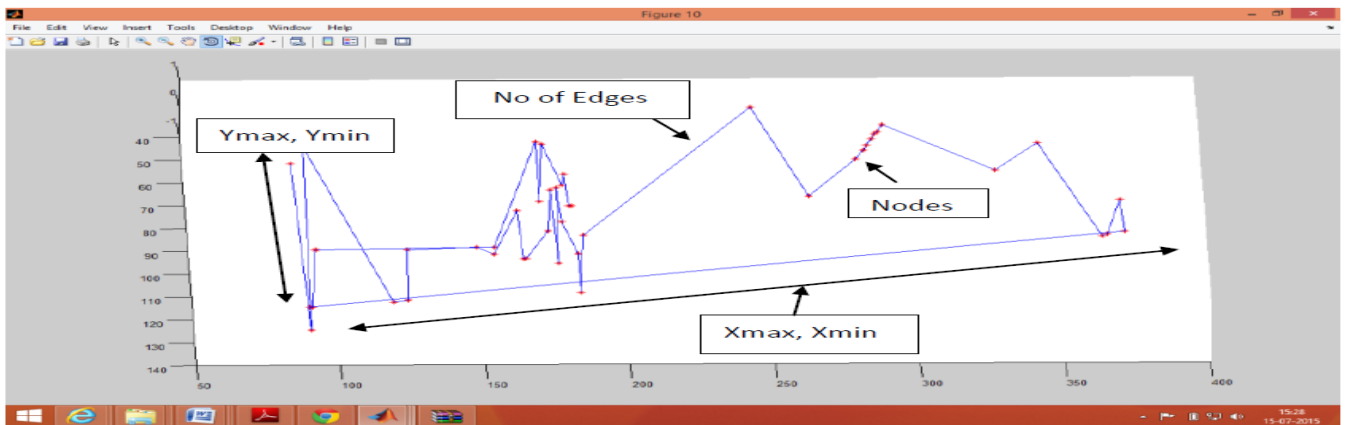


Figure: 6 X and Y co-ordinates, extreme points and number of edges, nodes and all points

VI. EXPERIMENTATION

6.1 Classification using euclidian distance

An algorithm that implements classification, especially in a concrete implementation, is known as a classifier. The term "classifier" sometimes also refers to the mathematical function, implemented by a classification algorithm that maps input data to a category. The classifiers are categorized as statistical methods, such as artificial neural networks, Euclidian distance, support vector machines and multiple classifier combination.

The Euclidean distance or Euclidean metric is the "ordinary" distance between two points that one would measure with a ruler, and is given by the Pythagorean formula. By using this formula as distance, Euclidean space (or even any inner product space) becomes a metric space. The associated norm is called the Euclidean norm.

The Euclidean distance between points p and q is the length of the line segment connecting them. In Cartesian coordinates, if $p = (p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$ are two points in Euclidean n -space, then the distance from p to q , or from q to p is given by equation 2. Using Euclidean distance metric, similarity score between any two feature set can be obtained in terms of the extracted features. The distance is used as the matching criterion, i.e. a signature is matched if this distance lies in a range of subjective threshold. However, using the Euclidean distance we generate matching scores by matching a query signature with all the signatures of database.

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

In order to evaluate designed system's performance, experiments have been carried out. The experiment has been performed on a signature database where 22 feature of each signature were extracted. The features extracted were graph based, which tend to be very efficient than other features such as local/global. The features extracted are listed as:

- X and Y co-ordinates of Extreme points(16)

- No of Points or Nodes(1)
- No of Edges(1)
- Minimum and maximum X-Co-ordinate value of among all the branch points (2)
- Minimum and maximum Y-Co-ordinate value of among all the branch points (2)

Identification rate resulted from this experiment is 94.4% performance. The performance was checked against number of persons. Initially experimented was started with database having 10 persons, from each person collected ten signatures. Gradually it has increased by 10 persons in each step. The different performances of system obtained with different sized database are shown in the figure 7.

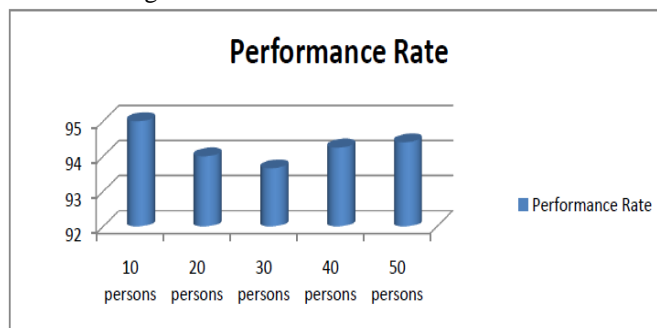


Figure: 7 Performance rate

The minimum performance 94.40% is obtained for the 50 persons, where 8 samples were used for each person during training and 2 samples completely different from training are taken for testing. The signature database comprises of an aggregate of 500 handwritten signature images, out of these, 472 signatures were recognised and 28 were not correctly recognized.

VII. CONCLUSION AND FUTURE WORK

Signature recognition system is an important behavioural biometric method and has wide range of applications. The proposed system has used extreme graph based parameters and Euclidean distance classifier to recognize signature. This system achieved the accuracy rate ranging from 94% to 95% for enrolment of 10-50 persons. There is an exponential improvement in identification rate using graph based technique and Euclidian distance matching parameter. But Euclidian distance takes more time for processing as the number of persons in the database increase. So far, it implements only some basic concepts of graph theory. Since this field is very wide, future work will be directed to find some more suitable graph features and combining with present work to increase the performance of handwritten signature identification system.

REFERENCES

- [1] Tomislav Fotak, Miroslav Baca, Petra Koruga, "Handwritten signature identification using basic concepts of graph theory", *Faculty of Organization and Informatics, Centre for biometrics University of Zagreb*, vol. 4,2011.
- [2] Vamsi Krishna Madasu,, "An Automatic Offline Signature Verification and Forgery Detection System", *pattern recognition applications and technology*, vol. 5,2008.
- [3] Meenakshi S Arya, A Preliminary Study on Various Off-line Hand Written Signature Verification Approaches, *International Journal of Computer Applications*, Volume 1 – No. 9,2010.
- [4] K. N. Pushpalatha , Aravind Kumar Gautham, D. R.Shashikumar, K. B. ShivaKumar and Rupam Das, "Offline Signature Verification with Random and Skilled Forgery Detection Using Polar Domain Features and Multi Stage Classification-Regression Model", *International Journal of Advanced Science and Technology*, Vol.59,2013.
- [5] Dipali K. Bhole1, A. V.Vidhate and Shrikant Velankar, "Offline Signature using Cross Validation and Graph Matching Approach", *International Journal of Technology And Engineering System(IJTES)*,Vol. 2,No1,2011.
- [6] K N Pushpalatha, Arvind Kumar Gautam, "HMM based Offline Signature Verification system using Contourlet Transform and Textural features", *International Journal of Engineering Science and Technology (IJEST)*, Vol. 6, No.7,2014.
- [7] M. Aretha and A. Govardhan, "An Efficient Offline Signature Verification System", *International Journal of Machine Learning and Computing*, Vol. 4, No. 6,2014.
- [8] Surabhi Garhawal, Neeraj Shukla, "A Study on Handwritten Signature Verification Approaches", *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, Volume 2, Issue 8,2013.
- [9] Ranjan Jana , Rituparna Saha Debaleena Datta, "Offline Signature Verification using Euclidian Distance", *International Journal of Computer Science and Information Technologies*, Vol. 5,2014.
- [10] Ashwini Pansare, Shalini Bhatia, "Off-line Signature Verification Using Neural Network," *International Journal of Scientific & Engineering Research*, Volume 3, 2012.