International Journal of Latest Engineering and Management Research (IJLEMR)
ISSN: 2455-4847
www.ijlemr.com || Volume 02 - Issue 08 || August 2017 || PP. 01-11

Digital Signature Representation Using SOM Method in OCR Technique

R. Ravi Chakravarthi
Lecturer in Information Technology Department,
Musanna College of Technology, Maladdah, P.O. Box: 191, PC- 314
Sultanate of Oman

Dr. E. Chandra
Professor & Head of Computer Science Department,
Bharathiar University, Coimbatore - 641046
Tamil Nadu. India

Abstract: Signature could be a means of writing one's own initials or name which may distinguish his/her identity from others and can be used for authentication functions. It's the most effective means of authenticating an individual since every individual possesses a different variety of writing. Two signatures will take issue from every other in terms of the pressure exerted whereas language, the shape of loops, the speed of writing, and numerous alternative options. Several algorithms are written to verify these signatures supported completely different sets of options extracted additionally as completely different classifiers used for classification. Optical Character Recognition (OCR) is the process of classification of optical patterns contained in a digital image. The character recognition is achieved through segmentation, feature extraction and classification. This chapter presents the basic ideas of OCR needed and the different techniques of OCR systems such as optical scanning, location segmentation, preprocessing, segmentation, representation, feature extraction, training and recognition, etc. The Self-organizing Feature Map (SOM) has been developed for the clustering of input vectors and has been commonly used as unsupervised learned classifiers. In this paper we describe the use of the SOM neural network model for signature verification. The triple-crown implementation of associate automatic biometric system depends primarily on the consistency of the used coaching sets. Signatures of an equivalent author square measure similar however not identical, since they'll disagree each globally and domestically, in location, scale, and orientation. In distinction with fingerprints, signatures that square measure fully authentic ne'er exist. This paper emphasizes the applying of a competitive neural specification for checking the consistency of the information set happiness to a private in an exceedingly biometric information. A neural network primarily based consistency live is projected to quantify the intra-variability of the people signatures. A replacement democratic neural specification is then given for step-down of the rejection error and maximization of the share of correct classification supported some well-known options and a replacement feature set.

Keywords:K-nearest neighbor's algorithm, Optical Character Recognition, Self Organizing Map, Image segmentation, feature extraction and classification.

I. Introduction

OCR belongs to the family of machine recognition techniques activity automatic identification. Automatic identification is that the method wherever the popularity system identifies objects mechanically, collects knowledge concerning them and enters knowledge directly into computer systems i.e. while not human intervention. The external knowledge is captured through analysis of pictures, sounds or videos. To capture knowledge, an electrical device is used that converts the particular image or sound into a digital file. The file is then kept and at a later time it is analyzed by the pc. Text detection and recognition in pictures and videos could be an analysis space that makes an attempt to develop an Automatic Data Processing (ADP) system with the power to mechanically scan from pictures and videos the text content visually embedded in advanced backgrounds.

As associate example of the final seeing problems, this ADP system, ought to answer 2 typical queries "Where & What": "where could be a text string?" and "what will the text string say?" in a picture or a video. In other words, victimization such a system, text embedded in advanced backgrounds is mechanically detected and every character or word is recognized. We tend to begin with a review of presently obtainable automatic identification techniques and outline OCR's position among them. The standard means of coming into knowledge in a very few Personal Computer is through the keyboard. However, this is often not continually the most effective or the foremost economical means. The automated identification might function an alternate in several cases. There exist varied techniques for automatic

identification. That cowls the wants for various application areas. Some notable technologies and their applications value mentioning excluding OCR area unit speech recognition, oftenness, vision systems, magnetic tape, bar code, ink and optical mark reading. These technologies are actively utilized in past decades [10]. Here we tend to introduce these technologies in short from application purpose of reading.

(a) In speech recognition systems spoken input from a predefined library of words area unit recognized. Such systems area unit speaker freelance and area unit typically used for reservations or telecommunication ordering of products. Another quite such systems area unit those that area unit accustomed acknowledge speaker instead of words for identification.

(b) The oftenness identification is commonly utilized in reference to toll roads for identification of cars. Special instrumentation is needed each to send and to scan the data. The data is inaccessible to humans.

(c) The vision systems area unit implemented through the usage camera wherever the objects area unit known by their form or size. This approach is usually utilized in automatons for recirculation of bottles. Sort the kind of bottle should be recognized initial because the quantity reimbursed for a bottle depends on its type.

(d) The information contained n magnetic stripes area unit wide used on credit cards, etc. quite a great amount of knowledge is kept on the magnetic tape however specially designed scanners area unit needed and therefore the information cannot be read by humans.

(e) The Universal Product Code consists of many dark and lightweight lines representing a code for associate eleven digit range, ten of that determine the actual product. The Universal Product Code is scanned optically once the merchandise moves over glass window by a centered ray of light of weak intensity that is swept across glass window in a very specially designed scanning pattern. The mirrored light-weight is measured and analyzed by computer. Thanks to early standardization bar codes area unit nowadays wide used and represent a serious share of the overall market place for automatic identification. The Universal Product Code represents a novel range that identifies the merchandise and a worth operation is important to retrieve information concerning the value. The binary pattern representing the Universal Product Code takes up a lot of area considering the tiny quantity of knowledge it really contains. The bar codes don't seem to be legible to humans.

(f) The printing in ink is especially used at intervals bank applications. The characters area unit written in ink contains finely ground magnetic material. They're written in artificial fonts that are specifically designed for the applying. Before the characters area unit browse the ink is exposed to a magnetic flux. This method accentuates every character and helps alter the detection. The characters area unit browse by deciphering the wave form obtained once scanning the characters horizontally. Every character is intended to possess its own specific wave form. Though designed for machine reading, the characters area unit still legible to humans. However, reading relies on characters being written with ink.

(g) The optical mark reading technology is employed to register location of marks. It's accustomed browse forms wherever the data is given by marking pre-defined alternatives. Such forms are clear to humans. This approach is economical once input is affected. It's predefined with fastened variety of alternatives.

OCR tries to deal with many problems with higher than mentioned techniques for automatic identification. They’re needed once the data is legible each to humans and machines. OCR systems have engraved a distinct segment place in pattern recognition. Their singularity lies within the undeniable fact that it doesn't need management of method that produces info. OCR deals with the matter of recognizing optically or printing has been completed whereas the net recognition is achieved where pc acknowledges the characters as they're drawn. Each hand written and written characters could also be recognized however the performance is directly dependent upon the standard of input documents. The lot of unnatural the input is, higher is that the performance of OCR system. However once it involves totals free handwriting performance of OCR machines remains questionable.

Personal identity verification is an important and frequently required application. There are several approaches taken using biometrics [5], [6], [8], [15]. Biometrics identify a person based on who he is rather than what he has (card, key, etc) or what he knows (password, PIN, etc). These can broadly be classified into physiological and behavioral biometrics. Physiological biometrics are based on the direct measurements of the physical parts (such as fingerprints, face, iris, hand geometry, etc) of the human body. Behavioral biometrics are based on the measurement of an action performed (such as signature, gait, speech, gesture, etc) by the individual [2]. A signature is perhaps the oldest and most widely accepted form of personal identity verification, especially in legal and financial matters. The characteristics of a signature can be divided into two types; i.e., static and dynamic. The static features are essentially the visible features such as the shape. Signature verification techniques based on these characteristics are referred to as offline signature verification. Since the static

characteristics are visible, it is possible for a skilled forger to reproduce these characteristics. On the other hand, dynamic features are those that are hidden, such as the pressure applied at the pen-tip, the position of the pen-tip, etc. Signature verification techniques based on these characteristics are referred to as online signature verification. Since the dynamic characteristics are not visible, it is not possible for any person to reproduce these characteristics as they incorporate the writer's personal style.

Automatic signature verification is essential due to the inability to identify signatures simply by visual inspection. Automatic inspection, though not perfect, is quite reliable and efficient. Considerable work has been done over the last few decades [9], [10], [12]. The last decade has seen an increased emphasis on the use of neural networks for this purpose [4], [14].

In this paper, we consider the recognition of an individual based on a behavioral biometric, specifically through his signature. We consider the dynamic characteristics of a signature and hence, we use online signature recognition.

1.1 Goals

The investigation of text detection and recognition in complicated background is actuated by leading edge applications of digital multimedia system. Nowadays additional audio and visual information is captured, stored, delivered and managed in digital forms. The wide usage of digital media files provokes several new challenges in mobile information acquisition and huge multimedia system management. Among the foremost outstanding are:

1. Automatic broadcast annotation: creates a structured, searchable read of archives of the broadcast content.
2. Digital media quality management: archives digital media files for economical media management.
3. Video piece of writing and cataloging: catalogs video databases on basis of content relevance;
4.Library digitizing: digitizes cowl of journals, magazines and varied videos victimization advanced image and video OCR.

II. Proposed Method

A typical OCR system consists of many elements as shown in Fig. 1 [3, 7].The primary step is to digitize Analog document victimization an optical scanner. Once regions containing text are settled every image is extracted through segmentation method. The extracted symbols are pre-processed, eliminating noise to facilitate feature extraction. The identity of every image is found by comparison extracted options with descriptions of image categories obtained through a previous learning section. Finally discourse data is employed to reconstruct words and numbers of the initial text. These steps are shortly conferred here. Interested readers will refer [11] for additional elaborate discussion of OCR system elements.

2.1 Optical Scanning

The first element in OCR is optical scanning. Through scanning method digital image of original document is captured. In OCR optical scanners are used that incorporate transport mechanism and sensing device that converts strength into gray levels. Written documents incorporate black print on white background. Once playing OCR structure image is regenerate into bi-level black and white image. This method referred to as thresholding is performed on scanner to avoid wasting memory area and process effort. The thresholding method is vital because the results of recognition are completely keen about quality of bi-level image. A hard and fast threshold is employed wherever grey levels below this threshold are black and levels on top of white. For top distinction document with uniform background a pre-chosen mounted threshold are often sufficient. However, documents encountered in follow have rather massive vary. In these cases, a lot of refined strategies for thresholding are needed to get smart results. The simplest thresholding strategies vary threshold adapting to native properties of document like distinction and brightness. However, such strategies sometimes rely upon structure scanning of document which needs a lot of memory and process capability. The flow diagram of the proposed method is shown below.

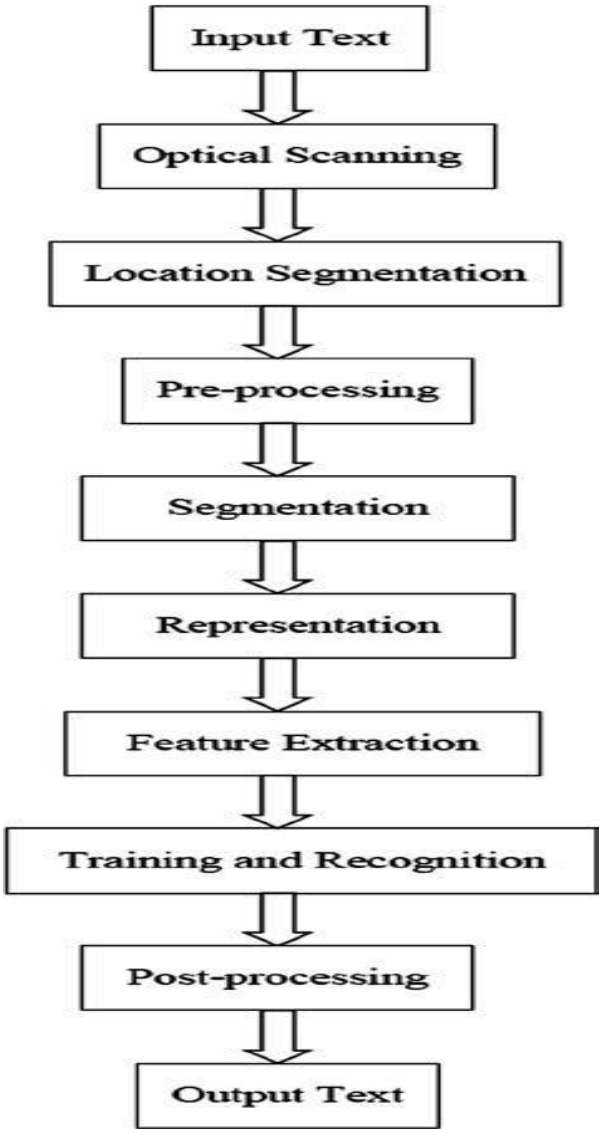

Figure 1: Flow Diagram of the proposed Method

2.2 LOCATION SEGMENTATION

The next OCR part is location segmentation. Segmentation determines constitutes of a picture. It's necessary to find regions of document that have written knowledge and are distinguished from figures and graphics. For instance, once activity automatic mail sorting through envelopes address should be set and separated from alternative prints like stamps and company logos, before recognition. Once applied to text, segmentation is isolation of characters or words. Most of OCR algorithms phase words into isolated characters that are recognized separately. Typically segmentation is performed by analytic every connected part. This system is straightforward to implement however issues arise if characters bit or they're fragmented and encompass many elements. The most issues in segmentation are: (a) extraction of touching and fragmented characters (b) distinctive noise from text (c) misinterpreting graphics and pure mathematics with text and the other way around. For interested readers additional details are obtainable in [11].

2.3 PREPROCESSING

The third OCR element is pre-processing. The data reckoning on the information acquisition kind is subjected to variety of preliminary process steps to create it usable within the descriptive stages of character analysis. The image ensuing from scanning method might contain certain quantity of noise. Reckoning on the

scanner resolution and therefore the inherent thresholding, the characters could also be dirty or broken. a number of these defects which can cause poor recognition rates and square measure eliminated through pre-processor by smoothing digitized characters. Smoothing implies each filling and dilution. Filling eliminates tiny breaks, gaps and holes in digitized characters whereas dilution reduces dimension of line. The foremost common technique for smoothing moves a window across binary image of character and applies sure rules to the contents of window. Pre-processing conjointly includes normalization in conjunction with smoothing. The normalization is applied to get characters of uniform size, slant and rotation. The right rotation is found through its angle. For turned pages and features of text, variants of Hough remodel square measure ordinarily used for police work skew.

The pre-processing component thus aims to produce data that are easy for the OCR systems to operate accurately. It is an important activity to be performed before the actual data analysis. The main objectives of pre-processing can be pointed as [1, 3]:
(a) Noise reduction (b) normalization of the data and (c) compression in the amount of information to be retained. In rest of this subsection the aforementioned objectives of pre-processing objectives are discussed with the corresponding techniques.

(a) **Noise reduction:** The noise introduced by the optical scanning device or the writing instrument causes disconnected line segments, bumps and gaps in lines, stuffed loops, etc. The distortion together with native variations, misreckoning of corners, dilation and erosion may be a potential drawback. It's necessary to eliminate these imperfections before actual process of the info. The noise reduction techniques are often classified in 3 major teams [1, 3]: (i) filtering (ii) morphological operations and (iii) noise modeling.
   (i) Filtering aims to get rid of noise and diminish spurious points typically introduced by uneven writing surface and poor rate of the information acquisition device. Varied abstraction and frequency domain filters have been designed for this purpose. The essential plan is to convolute a pre defined mask with the image to assign a price to a picture element as a perform of the grey values of its neighboring pixels. Many filters are designed for smoothing, sharpening, thresholding, removing slightly textured or colored background and distinction adjustment functions [1, 3].
   (ii) The basic plan behind the morphological operations is to filter the character image exchange the convolution operation by the logical operations. Numerous morphological operations are designed to attach the broken strokes, decompose the connected strokes, sleek the contours prune wild points, skinny the characters and extract the boundaries [1, 3]. The morphological operations is with success wont to remove noise on the character pictures as a result of quality of paper and ink moreover as erratic hand movement.
   (iii) Noise will usually be removed by activity techniques if it might have been doable to model it. However, noise modeling isn't possible in most of the applications. There exists some offered literature on noise modeling introduced by optical distortion like speckle, skew and blur. It's conjointly doable to assess the standard of the character pictures and remove the noise to a particular degree [1, 3].

(b) **Normalization:** The standardization ways aim to get rid of the variations of the writing and acquire standardized knowledge. A number of the unremarkably used ways for standardization square measure [1, 3]: (i) Skew Normalization and Baseline Extraction (ii) slant standardization (iii) size standardization and (iv) Contour smoothing.
   (i) Skew normalization and baseline extraction: attributable to inaccuracies within the scanning method and expressive style the writing could also be slightly leaning or arced at intervals the image. This may hurt the effectiveness of the algorithms and so ought to be detected and corrected. To boot, some characters distinguished in line with the relative position with relevance the baseline, the ways of baseline extraction embrace victimization the projection profile of the image, nearest neighbor cluster, cross correlation technique between lines and Hough rework [1, 3]. A pretty repulsive nearest neighbor is employed for extracting the baseline of sophisticated handwriting in significant noise [1, 3]. When skew detection the character or word is translated to the origin, turned or stretched till the baseline is horizontal and retranslated into the visual display unit area.
   (ii) Slant Standardization: One of the measurable factors of various hand-writing designs is that the slant angles between longest strokes in a word and therefore the vertical direction.

Slant standardization is employed to normalize all characters to a typical type. The foremost common methodology for slant estimation is that the calculation of the typical angle of close to vertical components. The vertical line components from contours area unit extracted by tracing chain code elements employing a combination of one dimensional filters [1, 3]. The coordinates of the beginning and finish points of every line element give the slant angle. The projection profiles area unit computed for variety of angles far from the vertical direction [1, 3]. The angle equivalent to the projection with the best positive spinoff is employed to notice quantity amount of overlap between vertical strokes and therefore the dominant slant angle. The slant detection is performed by dividing the image into vertical and horizontal windows [1, 3]. The slant is estimated supported the middle of gravity of the higher and lower 1/2 every window averaged over all the windows. A variant of the Hough trans-form is employed by scanning left to right across the image and calculative projections within the direction of twenty one totally different slants [1, 3]. The highest 3 projections for any slant area unit additional and therefore the slant with the most important count is taken because the slant price. In some cases the popularity systems don't use slant correction and compensate it throughout coaching stage [1, 3].

(iii) Size standardization is employed to regulate the character size to a precise customary. The OCR ways might apply for each horizontal and vertical size normalizations. The character is split into variety of zones and every of those zones is one by one scaled [1, 3]. The dimensions standardization also can be performed as a section of the coaching stage and therefore the size parameters are calculable one by one for every explicit coaching information [1, 3]. The sample characters are bit by bit shrunken to the optimum size that maximizes the popularity rate within the coaching information. The word recognition preserves giant intra category variations within the length of words so that they may additionally assist in recognition; it tends to solely involve vertical height standardization or bases the horizontal size standardization on the size issue calculated for vertical standardization [1, 3].

(iv) Contour smoothing eliminates the errors thanks to the erratic hand motion throughout the writing. It typically reduces the amount of sample points required to represent the script and so improves potency in remaining preprocessing steps [1, 3].

(c) **Compression:** It is well known that classical image compression techniques transform the image from the space domain to domains which are not suitable for recognition. The compression for OCR requires space domain techniques for preserving the shape information. The two popular compression techniques used are: (i) thresholding and (ii) thinning.

It is to be noted that the above techniques affect the data and may introduce unexpected distortions to the character image. As a result these techniques may cause the loss of important information about writing and thus should be applied with care.

2.4 SEGMENTATION

The preprocessing stage yields a clean character image within the sense that a sufficient quantity of form info, high compression, and low noise on a normalized image is obtained. Here the character image is meta meric into its subcomponents. Segmentation is very important as a result of the extent one will reach in separation of the assorted lines within the characters directly affects the popularity rate. Internal segmentation is employed here that isolates lines and curves within the cursively written characters. Although many remarkable strategies have developed within the past and a range of techniques have emerged, the segmentation of cursive characters is associate degree unsolved drawback. The character segmentation methods area unit divided into three classes [1, 3]: (a) express segmentation (b) implicit segmentation and (c) mixed methods.

2.4.1 Representation

The fifth OCR part is representation. The image representation plays one in all the foremost vital roles in any recognition system. Within the simplest case, grey level or binary pictures square measure fed to a recognizer. However, in most of the recognition systems so as to avoid further complexness and to extend the accuracy ofthe algorithms, a additional compact and characteristic illustration is needed. For this purpose, a group of options is extracted for every category that helps distinguish it

from different categories whereas remaining invariant to characteristic variations at intervals the category [1, 3]. The character image illustration strategies square measure usually categorized into 3 major groups: (a) international transformation and series enlargement (b) applied mathematics illustration and (c) geometrical and topological representation.

(a) Global transformation and series expansion: An eternal signal usually contains a lot of data than has to be pictured for the aim of classification. This might be true for distinct approximations of continuous signals also. A technique to represent a symbol is by a linear combination of a series of easier well outlined functions. The coefficients of the linear combination offer a compact secret writing called transformation or series growth. Deformations like translation and rotations square measure invariant below international transformation and series growth. Some common remodel and series growth strategies utilized in OCR are: (i) Fourier remodels (ii) Gabor remodel (iii) wavelets (iv) Moments and (v) Karhunen loeve growth. It is used in the National Institute of Standards and Technology (NIST) OCR system for form based handprint recognition [1, 3]. Since it requires computationally complex algorithms, the use of karhunen loeve features in OCR problems is not widespread. However, by the continuous increase of the computational power, it has become a realistic feature for the current OCR systems [1, 3].

(b) Statistical illustration: The representation of a personality image by organization of points takes care of favor variations to some extent. Though this kind of illustration doesn't enable the reconstruction of the first image, it's used for reducing the dimension of the feature set providing high speed and low complexness. A number of the main applied mathematics options used for character illustration are: (i) division (ii) crossings and distances and (iii) projections.

(c) Geometrical and topological representation: The various global and local properties of characters can be represented by geometrical and topological features with high tolerance to distortions and style variations. This type of representation may also encode some knowledge about the structure of the object or may provide some knowledge as to what sort of components make up that object. The topological and geometrical representations can be grouped into: (i) Extracting and counting topological structures (ii) Measuring and approximating the geometrical properties (iii) Coding and (iv) Graphs and trees.

In conclusion, the major goal of representation is to extract and select a set of features which maximizes the recognition rate with the least amount of elements. The feature extraction and selection is defined [1, 3] as extracting the most representative information from the raw data which minimizes the within class pattern variability while enhancing the between class pattern variability.

2.5 Feature Extraction

The sixth OCR element is feature extraction. The target of feature extraction is to capture essential characteristics of symbols. Feature extraction is accepted together of the foremost troublesome issues of pattern recognition. The foremost undemanding means of describing character is by actual formation image. Another approach is to extract sure options that characterize symbols however leaves the unimportant attributes. The techniques for extraction of such options area unit divided into 3 groups' viz. (a) distribution of points (b) transformations and series expansions and (c) structural analysis. The various teams of options area unit evaluated consistent with their noise sensitivity, deformation, easy implementation and use. The standards employed in this analysis are: (a) strength in terms of noise, distortions, vogue variation, translation and rotation and (b) sensible usage in terms of recognition speed, implementation quality and independence. A number of the usually used feature extraction techniques area unit model matching and correlation, transformations, distribution of points and structural analysis. For interested readers additional details area unit out there in [11].

Another necessary task related to feature extraction is classification. Classification is that the method of distinguishing every character and distribution to that correct character category. The two necessary classes of classification approaches for OCR area unit call a priori and structural strategies. In call a priori recognition character description is numerically delineated in feature vector. There may additionally be pattern characteristics derived from body of character that aren't as simply quantified. Here relationship between the characteristics might necessary once selecting category membership. as an example, if we all know that a personality consists of 1 vertical and one horizontal stroke it's going to be either 'L' or 'T'. The link between 2 strokes is needed to tell apart characters. The principal approaches to call a priori recognition area unit minimum distance classifiers, applied math classifiers and neural networks. In structural recognition syntactic strategies area unit the foremost prevailing approaches. an in depth discussion of those approaches is on the market in [3, 11].

2.6 Training and□ Recognition

The seventh OCR component is training and recognition. OCR systems extensively use the methodologies of pattern recognition which assigns an unknown sample into a predefined class. The OCR is investigated in four general approaches of pattern recognition as suggested in [1, 3]: (a) template matching (b) statistical techniques (c) structural techniques and (d) SOM. These approaches are neither necessarily independent nor disjointed from each other. Occasionally, an OCR technique in one approach can also be considered to be a member of other approaches. In all of the above approaches, OCR techniques use either holistic or analytic strategies for the training and recognition stages. The holistic strategy employs top down approaches for recognizing the full character eliminating the segmentation problem. The price for this computational saving is to constrain the problem of OCR to limited vocabulary. Also, due to the complexity introduced by the representation of a single character or stroke the recognition accuracy is decreased. On the other hand, the analytic strategies employ bottom up approach starting from stroke or character level and going toward producing a meaningful text. The explicit or implicit segmentation algorithms are required for this strategy, not only adding extra complexity to the problem but also introducing segmentation error to the system. However, with the cooperation of segmentation stage, the problem is reduced to the recognition of simple isolated characters or strokes, which can be handled for unlimited vocabulary with high recognition rates.

Among these methods SOM method gives a better accuracy when compared to the others. The reference module stores the data for training while the sample module consists of data for validation. In this paper, the intelligent decision module uses a neural network model; i.e, a Self organizing map (SOM). SOM is a data visualization technique which reduces the dimensions of data through the use of self-organizing neural networks. SOM reduces high dimensional data into a map of usually 1 or 2 dimensions by clustering similar data groups. This makes it easy to visualize and classify data. SOM neural network is created for the seven people using MATLAB. Four features were extracted from each of the three dynamic variables (p, x and y). Hence, each signature sample is represented by a pattern Fkl of twelve features given by

$$F_{kl} = \left[n_{kl}^{(x)} \; n_{kl}^{(y)} \; n_{kl}^{(p)} \; a_{kl}^{(x)} \; a_{kl}^{(y)} \; a_{kl}^{(p)} \; d_{kl}^{(x)} \; d_{kl}^{(y)} \; d_{kl}^{(p)} \; t_{kl}^{(x)} \; t_{kl}^{(y)} \; t_{kl}^{(p)} \right]^T$$

Where k represents the signer, l the sample number, x, y and p the dynamic characteristics and n, a, d and t the tokens features; i.e., length, average value, standard deviation and trend coefficient respectively. Each signer provides five samples for training; these are combined to form a 12 × 5 matrix Fk given by

$$F_k = \left[\; F_{k1} \quad F_{k2} \quad F_{k3} \quad F_{k4} \quad F_{k5} \; \right].$$

Finally, the data from the seven signers is combined into a single 12×35 matrix F, which is the reference data, given by

$$F = \left[\; F_1 \quad F_2 \quad F_3 \quad F_4 \quad F_5 \; \right].$$

The SOM created earlier needs to be trained using the reference data F. The duration of the training depends on the number of epochs. One epoch of training is defined as a single presentation of all input vectors to the network. The network is then updated according to the results of all those presentations. Training occurs until a maximum number of specified epochs have been reached or the performance goal is met. In this project, the number of epochs was determined by trial and error. As the number of epochs increased, the percentage error decreased. For more than 200 epochs, the improvement was not significant. Therefore, the training was stopped at 200 epochs.

III. Results and Discussions

To verify the algorithm proposed before, the algorithm was executed on the computer. All the programs were written with MATLAB language. Figure 2 (a) is the input image; Fig.2 (b) is the output image.

(a) Input Image (b) Resized image

(c) Thinning Image (d) Output Image

Figure 2: Step by step output for the proposed method

This section reports the results of the testing conducted on the SOM using data from the sample module. To evaluate a signature verification system, two types of error rates are usually defined: False Rejection Rate (FRR) or Error Type I, the percentage of rejection of genuine signatures and False Acceptance Rate (FAR) or Error Type II, the percentage of acceptance of forgeries. These errors for the SOM trained using 50 epochs and 200 epochs are shown in Table I.

TABLE I
SYSTEM PERFORMANCE.

User	Epochs = 50		Epochs = 200	
	FAR (%)	FRR (%)	FAR (%)	FRR (%)
1	0.00	8.33	0.00	8.33
2	15.00	10.00	10.00	0.00
3	20.00	20.00	20.00	20.00
4	0.00	60.00	0.00	26.67
5	20.00	6.67	20.00	6.67
6	25.00	12.50	15.00	0.00
7	0.00	60.00	0.00	53.33
Total	11.43	20.95	9.29	19.05

IV. Comparative Analysis

The comparative analysis can be represented through Chart.

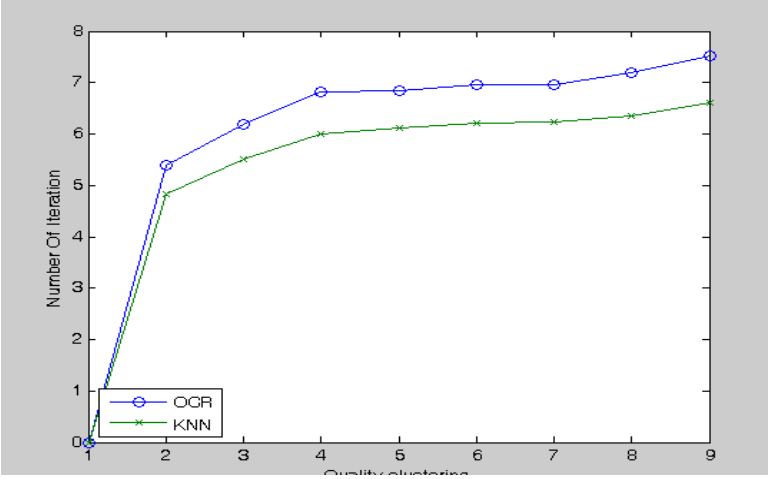

Figure 3: Comparative analysis of different methods

V. Conclusions

In this paper we have developed an online signature verification system using SOM and verified its performance *for Objectory design*. Three dynamic characteristics (*x*, *y*, and *p*) are measured and five features extracted from each. An SOM neural network is then trained and tested. The results are encouraging. The FAR and FRR can be improved further by increasing the reference sample size and/or also the number of features. However, a naive choice of the features may result in performance deterioration. In this case, the third central moment showed a high degree of sensitivity and system performance improved when it was discarded. The signature varies with the situation and condition under which it is obtained. Therefore, it is essential that the reference sample set is taken under different situations and conditions to capture all possibilities.

References

[1]. Arica, N., Vural, F. T. Y., An Overview of Character Recognition focused on Offline Handwriting, IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews, 31(2), pp 216–233, 2001.
[2]. Bunke, H., Wang, P. S. P. (Editors), Handbook of Character Recognition and Document Image Analysis, World Scientific, 1997.
[3]. Chaudhuri, A., Some Experiments on Optical Character Recognition Systems for dif-ferent Languages using Soft Computing Techniques, Technical Report, Birla Institute of Technology Mesra, Patna Campus, India, 2010.
[4]. Cheriet, M., Kharma, N., Liu, C. L., Suen, C. Y., Character Recognition Systems: A Guide for Students and Practitioners, John Wiley and Sons, 2007.
[5]. Dholakia, K., A Survey on Handwritten Character Recognition Techniques for various Indian Languages, International Journal of Computer Applications, 115(1), pp 17–21, 2015.
[6]. Mantas, J., An Overview of Character Recognition Methodologies, Pattern Recognition, 19(6), pp 425–430, 1986.
[7]. Rice, S. V., Nagy, G., Nartker, T. A., Optical Character Recognition: An Illustrated Guide to the Frontier, The Springer International Series in Engineering and Computer Science, Springer US, 1999.
[8]. Schantz, H. F., The History of OCR, Recognition Technology Users Association, Manchester Centre, VT, 1982.
[9]. Scurmann, J., Reading Machines, Proceedings of International Joint Conference on Pattern Recognition, Munich, pp 1031–1044, 1982.
[10]. Singh, S., Optical Character Recognition Techniques: A Survey, Journal of Emerging Trends in Computing and Information Sciences, 6 (4), pp 545–550, 2013.
[11]. Young, T. Y., Fu, K. S., Handbook of Pattern Recognition and Image Processing, Academic Press, 1986.
[12]. Yu, F. T. S., Jutamulia, S. (Editors), Optical Pattern Recognition, Cambridge University Press, 1998.
[13]. Kai Ding, Zhibin Liu, LianwenJin, Xinghua Zhu, A Comparative study of GABOR feature and gradient feature for handwritten 17hinese character recognition, International Conference on Wavelet Analysis and Pattern Recognition, pp. 1182-1186, Beijing, China, 2-4 Nov. 2007
[14]. Pranob K Charles, V.Harish, M.Swathi, CH.Deepthi, "A Review on the Various Techniques used for Optical Character Recognition", International Journal of Engineering Research and Applications, Vol.2, Issue 1, pp. 659-662, Jan-Feb 2012
[15]. Om Prakash Sharma, M. K. Ghose, Krishna Bikram Shah, "An Improved Zone Based Hybrid Feature Extraction Model for Handwritten Alphabets Recognition Using Euler Number", International Journal of Soft Computing and Engineering, Vol.2, Issue 2.
[16]. J. Pradeepa, E. Srinivasana, S. Himavathib,"Neural Network Based Recognition System Integrating Feature Extraction and Classification for English Handwritten",International journal of Engineering,Vol.25,No. 2, pp. 99-106, May 2012
[17]. Liu Cheng-Lin, Nakashima, Kazuki, H, Sako,H.Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, Pattern Recognition, Vol. 37, No. 2, pp. 265-279, 2004.
[18]. SupriyaDeshmukh, Leena Ragha, "Analysis of Directional Features - Stroke and Contour for Handwritten Character Recognition", IEEE International Advance Computing Conference, pp.1114-1118, 6-7 March, 2009,India
[19]. AmrithaSampath, Tripti C, Govindaru V, Freeman code based online handwritten character recognition for Malayalam using Back propagation neural networks, Advance computing: An international journal, Vol. 3,No. 4, pp. 51-58, July 2012.

[20]. RajibLochan Das, Binod Kumar Prasad, GoutamSanyal, "HMM based Offline Handwritten Writer Independent English Character Recognition using Global and Local Feature Extraction", International Journal of Computer Applications (0975 8887), Volume 46 No.10, pp. 45-50, May 2012.

[21]. Bhatia, N. and Vandana, Survey of Nearest Neighbor Techniques, International Journal of Computer Science and Information Security, Vol. 8, No. 2, (2001),302-305.

[22]. RajbalaTokas,ArunaBhadu, "A comparative analysis of feature extraction techniques for handwritten character recognition", International Journal of Advanced Technology & Engineering Research,Volume 2, Issue 4, pp. 215-219, July 2012.

[23]. AmrithaSampath, Tripti C, Govindaru V,"Freeman code based online handwritten character recognition for Malayalam using backpropagation neural networks",International journal on Advanced computing, Vol. 3, No. 4, pp. 51 - 58, July 2012.

[24]. Tirtharaj Dash, Time efficient approach to offline hand written character recognition using associative memory net.,International Journal of Computing and Business Research, Volume 3 Issue 3 September 2012.