An Overview Of Character Recognition Focused On Off-line Handwriting

Nafiz Arica, Student Member, IEEE and Fatos T. Yarman-Vural, Senior Member, IEEE

***Abstract*--Character Recognition (CR) has been extensively studied in the last half century and progressed to a level, sufficient to produce technology driven applications. Now, the rapidly growing computational power enables the implementation of the present CR methodologies and also creates an increasing demand on many emerging application domains, which require more advanced methodologies.**

This material serves as a guide and update for the readers, working in the Character Recognition area. First, an overview of CR systems and their evolution over time is presented. Then, the available CR techniques with their superiorities and weaknesses are reviewed. Finally, the current status of CR is discussed and directions for future research are suggested. Special attention is given to the off-line handwriting recognition, since this area requires more research to reach the ultimate goal of machine simulation of human reading.

***Index Terms*--Character Recognition, Off-line Handwriting Recognition, Segmentation, Feature Extraction, Training and Recognition.**

I. INTRODUCTION

Machine simulation of human functions has been a very challenging research field since the advent of digital computers. In some areas, which require certain amount of intelligence, such as number crunching or chess playing, tremendous improvements are achieved. On the other hand, humans still outperform even the most powerful computers in the relatively routine functions such as vision. Machine simulation of human reading is one of these areas, which has been the subject of intensive research for the last three decades, yet it is still far from the final frontier.

In this overview, Character Recognition (CR) is used as an umbrella term, which covers all types of machine recognition of characters in various application domains. The overview serves as an update for the state of the art in the CR field, emphasizing the methodologies required for the increasing needs in newly emerging areas, such as development of electronic libraries, multimedia databases and systems which require handwriting data entry. The study investigates the direction of the CR research, analyzing the limitations of methodologies for the systems, which can be classified based upon two major criteria: the data acquisition process (on-line or off-line) and the text type (machine-printed or hand-written). No matter which class the problem belongs, in general there are five major stages in the CR problem:

1. Pre-processing,
2. Segmentation,
3. Representation,
4. Training and recognition,
5. Post processing.

The paper is arranged to review the CR methodologies with respect to the stages of the CR systems, rather than surveying the complete solutions. Although the off-line and on-line character recognition techniques have different approaches, they share a lot of common problems and solutions. Since it is relatively more complex and requires more research compared to on-line and machine-printed recognition, off-line handwritten character recognition is selected as a focus of attention in this article. However, the article also reviews some of the methodologies for on-line character recognition, as it intersects with the off-line case.

After giving a historical review of the developments in Section 2, CR systems are classified in Section 3. Then, the methodologies of CR for pre-processing, segmentation, representation, training and recognition, post processing are reviewed in Section 4. Finally, the future research directions are discussed in Section 5. Since it is practically impossible to cite hundreds of independent studies conveyed in the field of CR, we suffice to provide only selective references and avoid an exhaustive list of studies, which can be reached from the references given at the end of this overview. The comprehensive survey on off-line and on-line handwriting recognition in [154], the survey in [179], dedicated to off-line cursive script recognition and the book in [133] which covers the Optical Character Recognition methodologies, can be taken as good starting points to reach the recent studies in various types and applications of the CR problem.

II. HISTORY

Writing, which has been the most natural mode of collecting, storing and transmitting the information through the centuries, now serves not only for the communication among humans, but also, serves for the communication of humans and machines. The intensive research effort on the field of CR was not only because of its challenge on simulation of human reading, but also, because it provides efficient applications such as the automatic processing of bulk amount of papers, transferring data into machines and web interface to paper documents. Historically, CR systems have evolved in three ages:

1900-1980 Early ages-- The history of character recognition can be traced as early as 1900, when the Russian

Manuscript received June 21, 1999.

N. Arica and F. T. Yarman-Vural are with the Computer Engineering Department, Middle East Technical University, Ankara, Turkey, (telephone: +90-312-2105595, e-mail: {nafiz, vural}@ ceng.metu.edu.tr).

Scientist Tyuring attempted to develop an aid for visually handicapped [123]. The first character recognizers appeared in the middle of the 1940s with the development of the digital computers [45]. The early work on the automatic recognition of characters has been concentrated either upon machine-printed text or upon small set of well-distinguished hand-written text or symbols. Machine-printed CR systems in this period generally used template matching in which an image is compared to a library of images. For handwritten text, low-level image processing techniques have been used on the binary image to extract feature vectors, which are then fed to statistical classifiers. Successful, but constrained algorithms have been implemented mostly for Latin characters and numerals. However, some studies on Japanese, Chinese, Hebrew, Indian, Cyrillic, Greek and Arabic characters and numerals in both machine-printed and handwritten cases were also initiated [134], [46], [135], [181].

The commercial character recognizers were available in 1950s, when electronic tablets capturing the x-y coordinate data of pen-tip movement was first introduced. This innovation enabled the researchers to work on the on-line handwriting recognition problem. A good source of references for on-line recognition until 1980 can be found in [180].

1980-1990 Developments-- The studies until 1980 suffered from the lack of powerful computer hardware and data acquisition devices. With the explosion on the information technology, the previously developed methodologies found a very fertile environment for rapid growth in many application areas, as well as CR system development [58], [19], [190]. Structural approaches were initiated in many systems in addition to the statistical methods. These systems broke the character image into a set of pattern primitives such as lines and curves. The rules were then determined which character most likely matched the extracted primitives [170], [189], [14]. However, the CR research was focused on basically the shape recognition techniques without using any semantic information. This led to an upper limit in the recognition rate, which was not sufficient in many practical applications. Historical review of CR research and development during this period can be found in [132] and [180] for off-line and on-line case, respectively.

After 1990 Advancements-- The real progress on CR systems is achieved during this period, using the new development tools and methodologies, which are empowered by the continuously growing information technologies.

In the early nineties, Image Processing and Pattern Recognition techniques are efficiently combined with the Artificial Intelligence methodologies. Researchers developed complex CR algorithms, which receive high-resolution input data and require extensive number crunching in the implementation phase. Nowadays, in addition to the more powerful computers and more accurate electronic equipments such as scanners, cameras and electronic tablets, we have efficient, modern use of methodologies such as Neural Networks, Hidden Markov Models, Fuzzy Set Reasoning and Natural Language Processing. The recent systems for the machine-printed off-line [13], [10] and limited vocabulary, user dependent on-line handwritten characters [125], [72], [152] are quite satisfactory for restricted applications. However, there is still a long way to go in order to reach the ultimate goal of machine simulation of fluent human reading, especially for unconstrained on-line and off-line handwriting.

III. CHARACTER RECOGNITION (CR) SYSTEMS

In this section, we classify the available CR systems according to the data acquisition techniques and the text type as follows:

A. Systems Classified According to the Data Acquisition Techniques

The progress in CR methodologies evolved in two categories according to the mode of data acquisition, as on-line and off-line character recognition systems.

The problem of recognizing handwriting, recorded with a digitizer, as a time sequence of pen coordinates is known as on-line character recognition. The digitizers are mostly electromagnetic-electrostatic tablets, which send the coordinates of the pen tip to the host computer at regular intervals. Some digitizers use pressure-sensitive tablets, which have layers of conductive and resistive material with a mechanical spacing between the layers. There are also, other technologies including laser beams and optical sensing of a light pen. The on-line handwriting recognition problem has a number of distinguishing features, which must be exploited to get more accurate results than the off-line recognition problem:

1. *It is a real time process.* While the digitizer captures the data during the writing, the CR system with or without a lag makes the recognition.
2. *It is adaptive in real time.* The writer gives immediate feedback to the recognizer for improving the recognition rate, as (s)he keeps drawing the symbols on the tablet and observes the results.
3. *It captures the temporal and dynamic information of the pen trajectory.* This information consists of the number and order of pen-strokes, the direction of the writing for each pen-stroke and the speed of the writing within each pen-stroke.
4. *Very little pre-processing is required.* The operations, such as smoothing, de-slanting, de-skewing, detection of line orientations, corners, loop and cusps are easier and faster with the pen trajectory data than on pixel images.
5. *Segmentation is easy.* Segmentation operations are facilitated by using temporal and pen-lift information, particularly, for hand-printed characters.

On the other hand, the disadvantages of the on-line character recognition are as follows:

1. The writer requires special equipment, which is not as comfortable as pen and paper.
2. It cannot be applied to documents printed or written on papers.
3. Punching is much faster and easier than handwriting for small size alphabet such as English.
4. The available systems are slow and recognition rates are low for handwriting that is not neat.

Applications of on-line character recognition systems include small hand-held devices, which call for a pen-only computer interface and complex multimedia systems, which use multiple input modalities including scanned documents, speech, keyboard and electronic pen. On-line character recognition systems are useful in social environments where speech does not provide enough privacy. They provide an efficient alternative for the large alphabets where the keyboard is cumbersome. Pen based computers [122], educational software for teaching handwriting [24] and signature verifiers [85] are the examples of popular tools utilizing the on-line character recognition techniques.

Off-line character recognition is known as Optical Character Recognition (OCR), because the image of writing is converted into bit pattern by an optically digitizing device such as optical scanner or camera. The recognition is done on this bit pattern data for machine-printed or hand-written text. The research and development is well progressed for the recognition of the machine-printed documents. In recent years, the focus of attention is shifted towards the recognition of hand-written script.

The major advantage of the off-line recognizers is to allow the previously written and printed texts to be processed and recognized. The drawbacks of the off-line recognizers, compared to on-line recognizers are summarized as follows:

1. Off-line conversion usually requires costly and imperfect pre-processing techniques prior to feature extraction and recognition stages.
2. The lack of temporal or dynamic information results in lower recognition rates compared to on-line recognition.

Some applications of the off-line recognition are large-scale data processing such as postal address reading [178], check sorting [94], office automation for text entry [56], automatic inspection and identification [164]. Off-line character recognition is a very important tool for creation of the electronic libraries. It provides a great compression and efficiency by converting the document image from any image file format into more useful formats like HTML or various word processor formats. Recently, content based image or video database systems make use of off-line character recognition for indexing and retrieval, extracting the writings in complex images [71]. Also, the wide spread use of web necessitates the utilization of off-line recognition systems for content based Internet access to paper documents [203].

B. Systems Classified According to the Text Type

Considering the text type, hand-written and machine-printed character recognition systems are two main areas of interest in the CR field.

Machine-printed text includes the materials such as books, newspapers, magazines, documents and various writing units in the video or still image. The problems for fixed-font, multi-font and omni-font character recognition is relatively well understood and solved with little constraint [190], [13], [10]. When the documents are generated on a high quality paper with modern printing technologies, the available systems yield as good as 99% recognition accuracy. However, the recognition rates of the commercially available products are very much dependent on the age of the documents, quality of paper and ink, which may result in significant data acquisition noise.

On the other hand, hand-written character recognition systems have still limited capabilities even for recognition of the Latin characters. The problem can be divided into two categories: cursive and hand-printed script. In practice, however, it is difficult to draw a clear distinction between them. A combination of these two forms can be seen frequently. Based on the nature of writing and the difficulty of segmentation process, Tappert [187] has defined five stages for the problem of handwritten word recognition as indicated in figure 1.

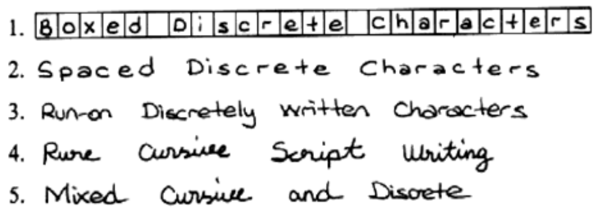


Fig. 1. Five stages of handwritten word recognition problem.

Boxed discrete characters require the writer to place each character within its own box on a form. The boxes themselves can be easily found and dropped out of the image or can be printed on the form in a special color ink that will not be picked up during scanning, thus eliminating the segmentation problem entirely. Spaced discrete characters can be segmented reliably by means of horizontal projections, creating a histogram of gray values in the image over all the columns and picking the valleys of this histogram as the points of segmentation. This has the same level of segmentation difficulty as is usually found with clean machine- printed documents. Characters at the third stage are usually discretely written, however they may be touching, therefore making the points of segmentation less obvious. Degraded machine-printed characters may also be found at this level of difficulty. There has been a fairly extensive and successful research on the first three stages (for overview and bibliography, see [58], [135], [196], [123]).

Cursively or mixed written texts require more sophisticated approaches compared to the previous cases. First of all, advanced segmentation techniques are to be used for character based recognition schemes. In pure cursive handwriting, a word is formed mostly from a single stroke. This makes segmentation by the traditional projection or connected-component methods ineffective. Secondly, shape discrimination between characters that look alike, such as U-V, I-1, O-0, is also difficult and requires the context information. In some languages, such as Arabic, there are characters, which can only be differed from the others according to the number or position of dots. A good source of references in hand-written character recognition can be found in [154], [4], [153], [133], [179].

IV. Methodologies of CR Systems

In this section, we focus on the methodologies of CR systems, emphasizing the off-line handwriting recognition problem. A hierarchical approach for most of the systems would be from pixel to text, as follows:

Pixel ⇒ Feature ⇒ Character ⇒ Sub –word ⇒ Word ⇒ Meaningful text

This bottom up approach varies a great deal, depending upon the type of the CR system and the methodology used. The literature review in the field of CR indicates that the above hierarchical tasks are grouped in the stages of the CR for pre-processing, segmentation, representation, training and recognition, post processing. In some methods, some of the stages are merged or omitted, in others a feedback mechanism is used to update the output of each stage. In this section, the available methodologies to develop the stages of the CR systems will be presented.

A. Pre-processing

The raw data, depending on the data acquisition type, is subjected to a number of preliminary processing steps to make it usable in the descriptive stages of character analysis. Pre-processing aims to produce data that are easy for the CR systems to operate accurately. The main objectives of pre-processing are:

1. Noise reduction,
2. Normalization of the data,
3. Compression in the amount of information to be retained.

In order to achieve the above objectives, the following techniques are used in the pre-processing stage:

1) Noise Reduction

The noise, introduced by the optical scanning device or the writing instrument, causes disconnected line segments, bumps and gaps in lines, filled loops etc. The distortion including local variations, rounding of corners, dilation and erosion, is also a problem. Prior to the character recognition, it is necessary to eliminate these imperfections. Hundreds of available noise reduction techniques can be categorized in three major groups [176], [168], as filtering, morphological operations and noise modeling.

Filtering: It aims to remove noise and diminish spurious points, usually introduced by uneven writing surface and/or poor sampling rate of the data acquisition device. Various spatial and frequency domain filters can be designed for this purpose. The basic idea is to convolute a pre-defined mask with the image to assign a value to a pixel as a function of the gray values of its neighboring pixels. Filters can be designed for smoothing [112], sharpening [113], thresholding [128], removing slightly textured or colored background [109] and contrast adjustment [155] purposes.

Morphological Operations: The basic idea behind the morphological operations is to filter the document image replacing the convolution operation by the logical operations. Various morphological operations can be designed to connect the broken strokes [8], decompose the connected strokes [29], smooth the contours, prune the wild points [168], thin the characters [160] and extract the boundaries [208]. Therefore, morphological operations can be successfully used to remove the noise on the document images due to low quality of paper and ink, as well as erratic hand movement.

Noise Modeling: Noise could be removed by some calibration techniques if a model for it were available. However, modeling the noise is not possible in most of the applications. There is very little work on modeling the noise introduced by optical distortion, like speckle, skew, and blur [11], [99], [151]. Nevertheless, it is possible to assess the quality of the documents and remove the noise to a certain degree, as suggested in [23].

2) Normalization

Normalization methods aim to remove the variations of the writing and obtain standardized data. The followings are the basic methods for normalization [59], [43].

Skew Normalization and Baseline Extraction: Due to inaccuracies in the scanning process and writing style, the writing may be slightly tilted or curved within the image. This can hurt the effectiveness of later algorithms and therefore should be detected and corrected. Additionally, some characters are distinguished regarding to the relative position with respect to the baseline (e.g. “9” and “g”). Methods of baseline extraction include using the projection profile of the image [83], a form of nearest neighbors clustering [68], cross correlation method between lines [34] and using the Hough transform [212]. Hough Transform is a technique for detecting curves by exploiting the duality between points on a curve and parameters of that curve. It is applied to characterize parameter curves of the writing [105]. In [145], attractive repulsive Neural Network is used for extracting the baseline of complicated handwriting in heavy noise (see figure 2). After skew detection, the character or word is translated to the origin, rotated or stretched until the baseline is horizontal and re-translated back into the display screen space.

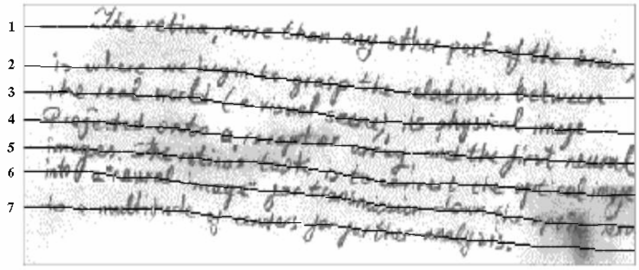


Fig. 2. Base line extraction using Attractive and Repulsive Network

Slant Normalization: One of the measurable factors of different handwriting styles is the slant angle between longest stroke in a word and the vertical direction. Slant normalization is used to normalize all characters to a standard form. The most commonly used method for slant estimation is the calculation of the average angle of near-vertical elements as in figure 3. In [119], vertical line elements from contours are extracted by tracing chain code components using a pair of one-dimensional filters. Coordinates of the start and end points of each line element provide the slant angle. Another study [60] uses an approach in which projection profiles are computed for a number of angles away from the vertical direction. The angle corresponding to the projection with the greatest positive derivative is used to detect the least amount of overlap between vertical strokes and therefore the dominant slant angle. In [19], slant detection is performed by dividing the image into vertical windows. These windows are then further divided horizontally by removing window portions that contain no writing. The slant is estimated based on the center of gravity of the upper and lower half of each window averaged over all the windows. Lastly, in [97] a variant of Hough transform is used by scanning left to right across the image and calculating projections in the direction of 21 different slants. The top three projections for any slant are added and the slant with the largest count is taken as the slant value. On the other hand, in some studies, recognition systems do not use slant correction and compensate it during training stage [7], [39].

Fig.3. Slant angle estimation: (a) near vertical elements, (b) average slant angle.

Size normalization: It is used to adjust the character size to a certain standard. Methods of character recognition may apply both horizontal and vertical size normalizations. In [209], the character is divided into number of zones and each of these zones is separately scaled. Size normalization can also be performed as a part of the training stage and the size parameters are estimated separately for each particular training data [6]. In Figure 4, two sample characters are gradually shrunk to the optimal size, which maximize the recognition rate in the training data. On the other hand, word recognition, due to the desire to preserve large intra-class differences in the length of words so they may assist in recognition, tends to only involve vertical height normalization or bases the horizontal size normalization on the scale factor calculated for the vertical normalization [97].

Contour Smoothing: It eliminates the errors due to the erratic hand motion during the writing. It, generally, reduces the number of sample points needed to represent the script, thus improves efficiency in remaining pre-processing steps [112], [8].

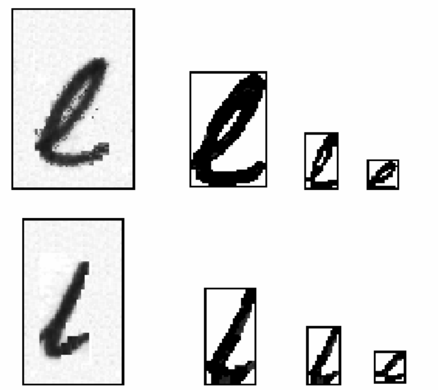


Fig.4. Normalization of characters "*e*" and "*l*" in [6].

3) 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. Compression for character recognition requires space domain techniques for preserving the shape information. Two popular compression techniques are thresholding and thinning.

Thresholding: In order to reduce storage requirements and to increase processing speed, it is often desirable to represent gray scale or color images as binary images by picking a threshold value Two categories of thresholding exist: *global* and *local*. Global thresholding picks one threshold value for the entire document image, often based on an estimation of the background level from the intensity histogram of the image [175]. Local (adaptive) thresholding use different values for each pixel according to the local area information [165]. In [191], a comparison of common global and local thresholding techniques is given by using an evaluation criterion that is goal-directed in the sense that the accuracies of a character recognition system using different techniques were compared. On those tested, it is shown that Niblack's locally adaptive method [138] produces the best result. Additionally, the recent study [207] develops an adaptive logical method by analyzing the clustering and connection characteristics of the characters in degraded document images.

Thinning: While it provides a tremendous reduction in data size, thinning extracts the shape information of the characters. Thinning can be considered as conversion of off-line handwriting to almost on-line like data, with spurious branches and artifacts. Two basic approaches for thinning are *pixel wise* and *non-pixel* wise thinning [104]. Pixel wise thinning methods locally and iteratively process the image until one pixel wide skeleton is remained. They are very sensitive to noise and may deform the shape of the character. On the other hand, the non-pixel wise methods use some global information about the character during the thinning. They produce a certain median or centerline of the pattern directly without examining all the individual pixels [12]. In [121], clustering based thinning method defines the skeleton of character as the cluster centers. Some thinning algorithms identify the singular points of the characters, such as end points, cross points and loops [210]. These points are the source of problems. In a non-pixel wise thinning, they are handled with global approaches. A survey of pixel wise and non-pixel wise thinning approaches is available in [104].

The iterations for thinning can be performed either in sequential or parallel algorithms. Sequential algorithms examine the contour points by raster scan [5] or contour following [57]. Parallel algorithms are superior to sequential ones, since they examine all the pixels simultaneously, using the same set of conditions for deletion [63]. They can be efficiently implemented in parallel hardware. An evaluation of parallel thinning algorithms for character recognition can be found in [103].

The pre-processing techniques are well explored and applied in many areas of image processing besides CR. The recent books (e.g. [27], [176]) on digital image processing serve as good sources of available techniques and references. Note that, the above techniques affect the data and may introduce unexpected distortions to the document image. As a result, these techniques may cause the loss of important information about writing. They should be applied with care.

B. Segmentation

The pre-processing stage yields a “clean” document in the sense that sufficient amount of shape information, high compression and low noise on normalized image is obtained. The next stage is segmenting the document into its sub components. Segmentation is an important stage, because the extent one can reach in separation of words, lines or characters directly affects the recognition rate of the script. There are two types of segmentation:

1. *External Segmentation*, which is the isolation of various writing units, such as paragraphs, sentences or words,
2. *Internal Segmentation*, which is the isolation of letters, specially, in cursively written words.

1) External Segmentation

External segmentation decomposes the page layout into its logical units. It is the most critical part of the document analysis, which is a necessary step prior to the off-line character recognition Although document analysis is a relatively different research area with its own methodologies and techniques, segmenting the document image into text and non-text regions is an integral part of the OCR software. Therefore, one who works in the CR field should have a general overview for document analysis techniques.

Page layout analysis is accomplished in two stages: The first stage is the *structural analysis*, which is concerned with the segmentation of the image into blocks of document components (paragraph, row, word, etc). The second one is the *functional analysis*, which uses location, size and various layout rules to label the functional content of document components (title, abstract, etc) [140].

A number of approaches regard a homogeneous region in a document image as a textured region. Page segmentation is then implemented by finding textured regions in gray-scale or color images. For example, Jain et al. use Gabor filtering and mask convolution [80], Tang et al.'s approach is based on fractal signature [184] and Doermann's method [42] employs wavelet multiscale analysis. Many approaches for page segmentation concentrate on processing background pixels or using the white space in a page to identify homogeneous regions [77]. These techniques include X-Y tree [28], pixel based projection profile [149], connected component based projection profile [65], white space tracing [2], and white space thinning [92]. They can be regarded as top-down approaches, which segment a page, recursively, by X-cut and Y-cut from large components, starting with the whole page to small components, eventually reaching individual characters. On the other hand, there is some bottom-up methods which recursively grow the homogeneous regions from small components based on the processing on pixels and connected components. An example of this approach may be Docstrum method, which uses k-nearest neighbor clustering [140]. Some techniques combine both top-down and bottom-up techniques [116]. A brief survey of the work in page decomposition can be found in [77].

2) Internal Segmentation

Internal Segmentation is an operation that seeks to decompose an image of a sequence of characters into sub-images of individual symbols. Although, the methods have developed remarkably in the last decade and a variety of techniques have emerged, segmentation of cursive script into letters is still an unsolved problem. Character segmentation strategies are divided into three categories [25].

Explicit Segmentation: In this strategy, the segments are identified based on “character like” properties. The process of cutting up the image into meaningful components is given a special name, “dissection”. Dissection is a process that analyzes an image without using specific class of shape information. The criterion for good segmentation is the agreement of general properties of the segments with those expected for valid characters. Available methods based on the dissection of an image use white space and pitch [70], vertical projection analysis [194], connected component analysis [197] and landmarks [67]. Moreover, explicit segmentation can be subjected to evaluation using linguistic context [107].

Implicit Segmentation-- This segmentation strategy is based on recognition. It searches the image for components that match predefined classes. Segmentation is performed by the use of recognition confidence, including syntactic or semantic correctness of the overall result. In this approach, two classes of methods can be employed; methods that make some search process and methods that segment a feature representation of the image [25].

The first class attempts to segment words into letters or other units without use of feature based dissection algorithms. Rather, the image is divided systematically into many overlapping pieces without regard to content. Conceptually, these methods originate from schemes developed for the recognition of machine-printed words [26]. The basic principle is to use a mobile window of variable width to provide sequences of tentative segmentations, which are confirmed by character recognition. Another technique combines dynamic programming and Neural Networks [22]. Lastly, the method of selective attention takes Neural Networks even further in the handling of the segmentation problem [49].

The second class of methods segments the image implicitly by classification of subsets of spatial features collected from the image as a whole. This approach can be divided into two categories; Hidden Markov Model Based approaches and Non-Markov based approaches. The survey in [55] provides an introduction to Hidden Markov Model based approaches in recognition applications. In [29] and [130], Hidden Markov Models are used to structure the entire word recognition process. Non-Markov approaches stem from concepts used in machine vision for recognition of occluded object [31]. This family of recognition-based approaches uses probabilistic relaxation [69], the concept of regularities and singularities [172] and backward matching [106].

Mixed Strategies: They combine explicit and implicit segmentation in a hybrid way. A dissection algorithm is applied to the image, but the intent is to “over segment”, i.e., to cut the image in sufficiently many places that the correct segmentation boundaries are included among the cuts made. Once this is assured, the optimal segmentation is sought by evaluation of subsets of the cuts made. Each subset implies a segmentation hypothesis, and classification is brought to bear to evaluate the different hypothesis and choose the most promising segmentation [47], [173]. In [108], the segmentation problem is formulated as finding the shortest path of a graph formed by binary and gray level document image. In [7], the HMM probabilities, obtained from the characters of a dissection algorithm, is used to form a graph. The optimum path of this graph improves the result of the segmentation by dissection and HMM recognition. Figure 5.a indicates the initial segmentation intervals, obtained by evaluating the local maxima and minima together with the slant angle information. Figure 5.b and c show the shortest path for each segmentation interval and the resulting candidate characters, respectively. Mixed strategies yield better results compared to explicit and implicit segmentation methods.

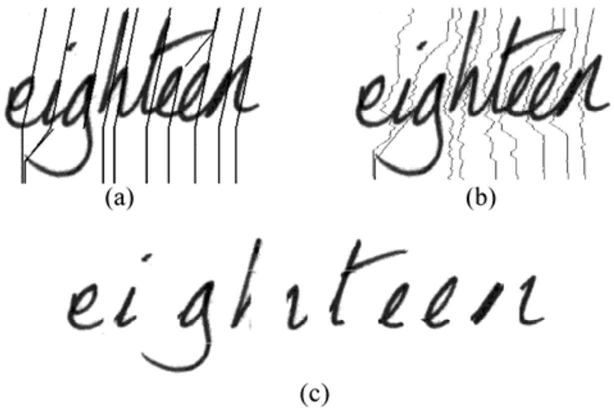


Fig.5. Segmentation by finding the shortest path of a graph formed by gray level image (a) Segmentation intervals, (b) Segmentation paths and (c) Segments.

The techniques presented above have limited capabilities in segmentation. Error detection and correction mechanisms should be embedded into the systems for which they were developed. As Casey and Lecolinet [25] pointed out, the wise use of context and classifier confidence leads to improved accuracy.

C. Representation

Image representation plays one of the most important roles in a recognition system. In the simplest case, gray-level or binary images are fed to a recognizer. However, in most of the recognition systems, in order to avoid extra complexity and to increase the accuracy of the algorithms, a more compact and characteristic representation is required. For this purpose, a set of features is extracted for each class that helps distinguish it from other classes, while remaining invariant to characteristic differences within the class [142]. A good survey on feature extraction methods for character recognition can be found in [192]. In the following, hundreds of document image representation methods are categorized in three major groups:

1) Global Transformation and Series Expansion

A continuous signal generally contains more information than needs to be represented for the purpose of classification. This may be true for discrete approximations of continuous signals as well. One way to represent a signal is by a linear combination of a series of simpler well-defined functions. The coefficients of the linear combination provide a compact encoding known as transformation or/and series expansion. Deformations like translation and rotations are invariant under global transformation and series expansion. Common transform and series expansion methods used in the CR field are:

Fourier Transforms: The general procedure is to choose magnitude spectrum of the measurement vector as the features in an n-dimensional Euclidean space. One of the most attractive properties of the Fourier Transform is the ability to recognize the position-shifted characters, when it observes the magnitude spectrum and ignores the phase. Fourier Transforms has been applied to CR in many ways [215], [199].

Gabor Transform: It is a variation of the windowed Fourier Transform. In this case, the window used is not a discrete size, but is defined by a Gaussian function [66].

Wavelets: Wavelet transformation is a series expansion technique that allows us to represent the signal at different levels of resolution. The segments of document image, which may correspond to letters or words, are represented by wavelet coefficients, corresponding to various levels of resolution. These coefficients are then fed to a classifier for recognition [111], [169]. The representation in multiresolution analysis (MRA) with low resolution can absorb the local variation in handwriting than in MRA with high resolution. However, the representation in low resolution may cause the important details for the recognition stage to be lost.

Moments: Moments, such as central moments, Legendre moments, Zernike moments, form a compact representation of the original document image that make the process of recognizing an object scale, translation, and rotation invariant [89], [37]. Moments are considered as series expansion representation, since the original image can be completely reconstructed from the moment coefficients.

Karhunen-Loeve Expansion: It is an eigen-vector analysis, which attempts to reduce the dimension of the feature set by creating new features that are linear combinations of the original ones. It is the only optimal transform in terms of information compression. Karhunen-Loeve expansion is used in several pattern recognition problems such as face recognition. It is also used in the National Institute of Standards and Technology (NIST) OCR system for form-based handprint recognition [53]. Since it requires computationally complex algorithms, the use of Karhunen-Loeve features in CR problems is not widespread. However, by the increase of the computational power, it will become a realistic feature for CR systems in the next few years [192].

2) Statistical Representation

Representation of a document image by statistical distribution of points takes care of style variations to some extent. Although this type of representation does not allow the reconstruction of the original image, it is used for reducing the dimension of the feature set providing high speed and low complexity. The followings are the major statistical features used for character representation:

Zoning: The frame containing the character is divided into several overlapping or non-overlapping zones. The densities of the points or some features in different regions are analyzed and form the representation. For example, contour direction features measure the direction of the contour of the character [131], which are generated by dividing the image array into rectangular and diagonal zones and computing histograms of chain codes in these zones. Another example is the bending point features, which represent high curvature points, terminal points and fork points [183]. Figure 6 indicates contour direction and bending point features.

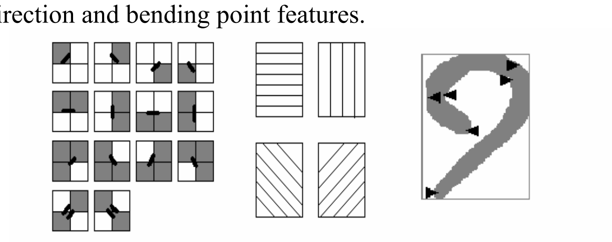


Fig.6. Contour direction and bending point features with zoning [131].

Crossings and Distances: A popular statistical feature is the number of crossing of a contour by a line segment in a specified direction. In [6], the character frame is partitioned into a set of regions in various directions and then the black runs in each region are coded by the powers of two. Another study [130] encodes the location and number of transitions from background to foreground pixels along vertical lines through the word. Also, the distance of line segments from a given boundary, such as the upper and lower portion of the frame, can be used as statistical features [20]. These features imply that a horizontal threshold is established above, below and through the center of the normalized script. The number of times the script crosses a threshold becomes the value of that feature. The obvious intent is to catch the ascending and descending portions of the script.

Projections: Characters can be represented by projecting the pixel gray values onto lines in various directions. This representation creates one-dimensional signal from a two dimensional image, which can be used to represent the character image [200], [186].

3) Geometrical and Topological Representation

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. Hundreds of topological and geometrical representations can be grouped in four categories:

Extracting and Counting Topological Structures: In this group of representation a pre-defined structure is searched in a character or word. The number or relative position of these structures within the character forms a descriptive representation. Common primitive structures are the strokes, which make up a character. These primitives can be as simple as lines (l) and arcs (c) which are the main strokes of Latin characters and can be as complex as curves and splines making up Arabic or Chinese characters. In on-line character recognition, a stroke is also defined as a line segment from pen-down to pen-up [180]. Characters and words can be successfully represented by extracting and counting many topological features such as, the extreme points, maxima and minima, cusps above and below a threshold, openings to the right, left, up and down, cross (x) points, branch (T) points, line ends (J), loops, direction of a stroke from a special point, inflection between two points, isolated dots, a bend between two points, symmetry of character, horizontal curves at top or bottom, straight strokes between two points, ascending, descending and middle strokes and relations among the stroke that make up a character, etc [21], [139], [120], [62], [119]. Figure 7 indicates some of the topological features.

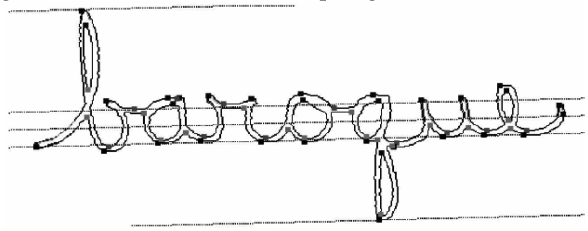


Fig.7. Topological features: maxima and minima on the exterior and interior contours, reference lines, ascenders and descenders [119].

Measuring and Approximating the Geometrical Properties: In many studies (e.g. [101], [39]), the characters are represented by the measurement of the geometrical quantities such as, the ratio between width and height of the bounding box of a character, the relative distance between the last point and the last y-min, the relative horizontal and vertical distances between first and last points, distance between two points, comparative lengths between two strokes, width of a stroke, upper and lower masses of words, word length. A very important characteristic measure is the curvature or change in the curvature [144]. Among many methods for measuring the curvature information one is suggested by [141] for measuring the directional distance and measures local stroke direction distribution for directional decomposition of the character image.

The measured geometrical quantities can be approximated by a more convenient and compact geometrical set of features. A class of methods includes polygonal approximation of a thinned character [162]. A more precise and expensive version of the polygonal approximation is cubic spline representation [127].

Coding: One of the most popular coding schema is Freeman's chain code. This coding is essentially obtained by mapping the strokes of a character into a 2-dimensional parameter space, which is made up of codes as shown in figure 8. There are many versions of chain coding. As an example, in [61], the character frame is divided to left-right sliding window and each region is coded by the chain code.

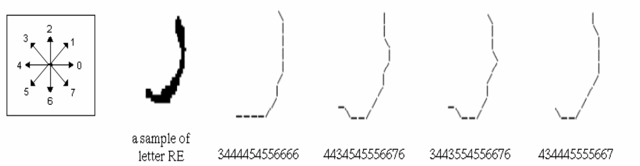


Fig.8. A sample Arabic character and the chain codes of its skeleton.

Graphs and Trees: Words or characters are first partitioned into a set of topological primitives, such as strokes, loops, cross points etc. Then, these primitives are represented using attributed or relational graphs [114]. There are two kinds of image representation by graphs. The first kind uses the coordinates of the character shape [36], [166]. The second kind is an abstract representation with nodes corresponding to the strokes and edges corresponding to the relationships between the strokes [117]. Trees can also be used to represent the words or characters with a set of features, which has a hierarchical relation [118].

The feature extraction process is performed mostly on binary images. However, binarization of a gray level image may remove important topological information from characters. In order to avoid this problem, some studies attempt to extract features directly from gray scale character images [110].

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. In [93], feature extraction and selection is defined as extracting the most representative information from the raw data, which minimizes the within class pattern variability while enhancing the between class pattern variability.

Feature selection can be formulated as a Dynamic Programming problem for selecting the k-best features out of N features, with respect to a cost function such as Fishers Discrimant ratio. Feature selection can also be accomplished by using Principal Component Analysis or a Neural Network trainer. In [38], the performance of several feature selection methods for character recognition are discussed and compared. Selection of features using a methodology as mentioned here requires expensive computational power and most of the time yields a sub optimal solution [51]. Therefore, the feature selection is, mostly, done by heuristics or by intuition for a specific type of the CR application.

D. Training and Recognition Techniques

CR systems extensively use the methodologies of pattern recognition, which assigns an unknown sample into a pre-defined class. Numerous techniques for CR can be investigated in four general approaches of Pattern Recognition, as suggested in [81]:

1. Template Matching,
2. Statistical Techniques,
3. Structural Techniques,
4. Neural Networks.

The above approaches are neither necessarily independent nor disjoint from each other. Occasionally, a CR technique in one approach can also be considered to be a member of other approaches.

In all of the above approaches, CR techniques use either holistic or analytic strategies for the training and recognition stages: Holistic strategy employs top down approaches for recognizing the full word, eliminating the segmentation problem. The price for this computational saving is to constrain the problem of CR to limited vocabulary. Also, due to the complexity introduced by the representation of whole cursive word (compared to the complexity of a single character or stroke), the recognition accuracy is decreased.

On the other hand, the analytic strategies employ bottom up approaches starting from stroke or character level and going towards producing a meaningful text. 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 (see table I).

TABLE I
STRATEGIES OF THE CR

Holistic Strategy	Analytic Strategy
Whole Word Recognition	Sub-word or Letter Recognition
Limited Vocabulary	Unlimited Vocabulary
Vulnerable to Recognition of Long Words	Vulnerable to Segmentation Errors
No Segmentation	Requires Explicit or Implicit Segmentation

1) Template Matching

CR techniques vary widely according to the feature set selected from the long list of features, described in the previous section for image representation. Features can be as simple as the gray-level image frames with individual characters or words or as complicated as graph representation of character primitives. The simplest way of character recognition is based on matching the stored prototypes against the character or word to be recognized. Generally speaking, matching operation determines the degree of similarity between two vectors (group of pixels, shapes, curvature etc.) in the feature space. Matching techniques can be studied in three classes:

Direct Matching: A gray-level or binary input character is directly compared to a standard set of stored prototypes. According to a similarity measure (e.g.: Euclidean, Mahalanobis, Jaccard or Yule similarity measures etc.), a prototype matching is done for recognition. The matching techniques can be as simple as one-to-one comparison or as complex as decision tree analysis in which only selected pixels are tested. A template matcher can combine multiple information sources, including match strength and k-nearest neighbor measurements from different metrics [195], [50]. Although direct matching method is intuitive and has a solid mathematical background, the recognition rate of this method is very sensitive to noise.

Deformable Templates and Elastic Matching: An alternative method is the use of deformable templates, where an image deformation is used to match an unknown image against a database of known images. In [78], two characters are matched by deforming the contour of one, to fit the edge strengths of the other. A dissimilarity measure is derived from the amount of deformation needed, the goodness of fit of the edges and the interior overlap between the deformed shapes (see figure 9).

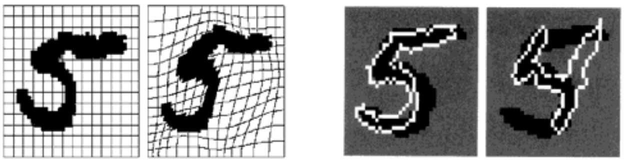


Fig.9. (a) Deformations of a sample digit, (b) Deformed template superimposed on target image, with dissimilarity measures in [78].

The basic idea of elastic matching is to optimally match the unknown symbol against all possible elastic stretching and compression of each prototype. Once the feature space is formed, the unknown vector is matched using dynamic programming and a warping function [73], [188]. Since the curves obtained from the skeletonization of the characters could be distorted, elastic matching methods cannot deal with topological correlation between two patterns in the off-line CR. In order to avoid this difficulty, a self-organization matching approach is proposed in [115] for hand-printed character recognition, using thick strokes. Elastic matching is also popular in on-line recognition systems [137].

Relaxation Matching: It is a symbolic level image matching technique that uses feature-based description for the character image. First, the matching regions are identified. Then, based on some well-defined ratings of the assignments, the image elements are compared to the model. This procedure requires a search technique in a multi-dimensional space, for finding the global maximum of some functions [157], [206]. Huang et al. proposed a multi font Chinese character recognition system in [98], where sampling points, including cross, branch and end points on the skeleton are taken as nodes of a graph. Each character class is represented by a constrained graph model, which captures the geometrical and topological invariance for the same class. Recognition is then made by a relaxation matching algorithm. In [204], Xie et al. proposed a handwritten Chinese character system, where small number of critical structural features, such as end points, hooks, T-shape, cross and corner are used. Recognition is done by computing the matching probabilities between two features by a relaxation method.

The matching techniques mentioned above are sometimes used individually or combined in many ways as part of the CR schemes.

2) Statistical Techniques

Statistical decision theory is concerned with statistical decision functions and a set of optimality criteria, which maximizes the probability of the observed pattern given the model of a certain class [41]. Statistical techniques are, mostly, based on three major assumptions:

1. Distribution of the feature set is Gaussian or in the worst case uniform,
2. There are sufficient statistics available for each class,
3. Given ensemble of images $\{I\}$, one is able to extract a set of features $\{f_i\} \in F$, $i \in \{1,...,n\}$, which represents each distinct class of patterns.

The measurements taken from n-features of each word unit can be thought to represent an n-dimensional vector space and the vector, whose coordinates correspond to the measurements taken, represents the original word unit. The major statistical approaches, applied in the CR field are the followings:

Non-parametric Recognition: This method is used to separate different pattern classes along hyper planes defined in a given hyperspace. The best known method of non-parametric classification is the Nearest Neighbor (NN) and is extensively used in CR [174]. It does not require a priori information about the data. An incoming pattern is classified using the cluster, whose center is the minimum distance from the pattern over all the clusters.

Parametric Recognition: Since a priori information is available about the characters in the training data, it is possible to obtain a parametric model for each character [15]. Once the parameters of the model, which is based on some probabilities, are obtained, the characters are classified according to some decision rules such as maximum Likelihood or Bayes method.

Clustering Analysis: The clusters of character features, which represent distinct classes, are analyzed by way of clustering methods. Clustering can be performed either by agglomerative or divisive algorithms. The agglomerative algorithms operate step-by-step merging of small clusters into larger ones by a distance criterion. On the other hand, the divisive methods split the character classes under certain rules for identifying the underlying character [211].

Hidden Markov Modeling (HMM): Hidden Markov Models are the most widely and successfully used technique for handwritten character recognition problem [129], [130], [96], [29], [30]. It is defined as a stochastic process generated by two interrelated mechanisms; a Markov Chain having a finite number of states and a set of random functions, each of which is associated with a state [158]. At discrete instants of time, the process is assumed to be in some state and an observation is generated by the random function corresponding to the current state. The underlying Markov chain then changes states according to its transitional probabilities. Here, the job is to build a model that explains and characterizes the occurrence of the observed symbols [82]. The output corresponding to a single symbol can be characterized as discrete or continuous. Discrete outputs may be characters from a finite alphabet or quantized vectors from a codebook, while continuous outputs are represented by samples from a continuous waveform. In generating a word or a character, the system passes from one state to another, each state emitting an output according to some probabilities until the entire word or character is out. There are two basic approaches to CR systems using HMM:

1. *Model Discriminant HMM*: A model is constructed for each class (word, character or segmentation unit), in the training phase. States represent cluster centers for the feature space. The goal of classification is then to decide on the model, which produces the unknown observation sequence. In [54] and [130], each column of the word image is represented as a feature vector by labeling each pixel according to its location. A separate model is trained for each character using word images, where the character boundaries have been identified. Then, the word matching process uses models for each word in a supplied lexicon. The advantage of this technique is that the word needs not be segmented into

characters for the matching process. In segmentation-based approaches, the segmentation process is applied to the word image and several segmentation alternatives are proposed [7], [96]. Each segmentation alternative is embedded to the HMM recognizer which assigns a probability for each segment. In post-processing, a search algorithm takes into account of the segmentation alternatives and their recognition scores in order to find the final result. Model discriminant HMM's also used for isolated character recognition in various studies [97], [6]. Figure 10 shows a model discriminant HMM with five states left-to-right topology, constructed for each character class.

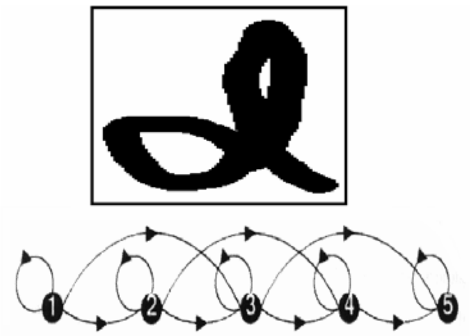


Fig.10. Model discriminant HMM with 5 states for the sample character "d".

2. *Path Discriminant HMM*: In this approach, a single HMM is constructed for the whole language or context. Modeling is supported by the initial and transitional probabilities on the basis of observations from a random experiment, generally, by using a lexicon. Each state may signify a complete character, a partial character or joint characters. Recognition consists of estimation of the optimal path for each class using Viterbi algorithm, based on dynamic programming. In [29] and [30], the input word image is first segmented into a sequence of segments in which an individual segment corresponds to a state. The HMM parameters are estimated from the lexicon and the segmentation statistics of the training images. Then, a modified Viterbi algorithm is used to find the best L-state sequences in recognition of a word image (see figure 11).

The performances of these two approaches are compared in various experiments by utilizing different lexicon sizes in [100]. The major design issue in the HMM problem is the selection of the feature set and HMM topology. These two tasks are strongly related to each other and there is no systematic approach, developed for this purpose.

Fuzzy Set Reasoning: Instead of using a probabilistic approach, this technique employs fuzzy set elements in describing the similarities between the features of the characters. Fuzzy set elements give more realistic results, when there is no a priori knowledge about the data and therefore the probabilities cannot be calculated. The characters can be viewed as a collection of strokes, which is compared to reference patterns by fuzzy similarity measures. Since the strokes under consideration are fuzzy in nature, the concept of fuzziness is utilized in the similarity measure. In order to recognize a character, an unknown input character is matched with all the reference characters and is assigned to the class of the reference character with the highest score of similarity among all the reference characters. In [35], fuzzy similarity measure is utilized to define Fuzzy Entropy for off-line handwritten Chinese characters. An off-line handwritten character recognition system is proposed in [1] using a fuzzy graph theoretic approach, where each character is described by a fuzzy graph. A fuzzy graph-matching algorithm is, then, used for recognition. Wang and Mendel propose an off-line handwritten recognition algorithm system, which generates crisp features first, and then, fuzzify the characters by rotating or deforming them [198]. The algorithm uses average values of membership for final decision. In handwritten word recognition, Gader et al. uses the choquet fuzzy integral as the matching function [52]. In on-line handwriting recognition, Plamondon proposes a fuzzy-syntactic approach to allograph modeling [146].

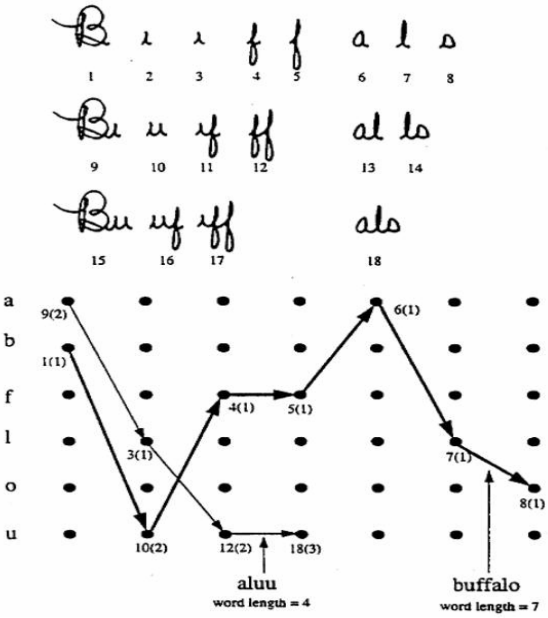


Fig.11. Path discriminant HMM: all block images of up to three segments (top) and two particular paths, or state sequences [29].

3) *Structural Techniques*

The recursive description of a complex pattern in terms of simpler patterns based on the shape of the object was the initial idea behind the creation of structural pattern recognition. These patterns are used to describe and classify the characters in the CR systems. The characters are represented as the union of the structural primitives. It is assumed that the character primitives extracted from writing are quantifiable and one can find the relations among them. The following structural methods are applied to the CR problems:

Grammatical Methods: In mid 1960's, researchers started to consider the rules of linguistics for analyzing the speech and writing. Later, various orthographic, lexicographic and linguistic rules were applied to the recognition schemes. The grammatical methods create some production rules in order to form the characters from a set of primitives through formal grammars. These methods may combine any type of topological and statistical features under some syntactic and/or semantic rules, [189], [170], [148]. Formal tools, like language theory, allow us to describe the admissible constructions and to extract the contextual information about the writing by using various types of grammars, such as string grammars, graph grammars, stochastic grammars and picture description language [193].

In grammatical methods, training is done by describing each character by a grammar G_i. In the recognition phase, the string, tree or graph of any writing unit (character, word or sentence) is analyzed in order to decide to which pattern grammar it belongs [14]. Top-down or bottom-up parsing does syntax analysis. Given a sentence, a derivation of the sentence is constructed and the corresponding derivation tree is obtained. The grammatical methods in the CR area are applied in character [173], word [87] and sentence [177] levels. In character level, Picture Description Language (PDL) is used to model each character in terms of a set of strokes and their relationship. This approach is used for Indian character recognition, where Devanagari characters are presented by a PDL [173]. The system stores the structural description in terms of primitives and the relations. Recognition involves a search for the unknown character, based on the stored description. In word level, bi-gram and three-gram statistics are used to form word generation grammars. Word and sentence representation uses knowledge bases with linguistic rules. Grammatical methods are mostly used in the post processing stage for correcting the recognition errors [18], [167].

Graphical Methods: Writing units are represented by trees, graphs, di-graphs or attributed graphs. The character primitives (e.g. strokes) are selected by a structural approach, irrespective of how the final decision making is made in the recognition [88], [182], [201]. For each class, a graph or tree is formed in the training stage to represent strokes, letters or words. Recognition stage assigns the unknown graph to one of the classes by using a graph similarity measure.

There are a great variety of approaches that use the graphical methods. Hierarchical graph representation approach is used for handwritten Korean and Chinese character recognition in [88] and [117] respectively. Pavlidis and Rocha use homoemorphic subgraph matching method for complete word recognition in [163]. First, the word image is converted into a feature graph. Edges of the feature graph are the skeletons of the input strokes. Graph nodes correspond to singularities (inflection points, branch points, sharp corners and ending points) on the strokes. Then, the meaningful subgraphs of the feature graph are recognized, matching the previously defined character prototypes. Finally, each recognized subgraph is introduced as a node in a directed net that compiles different interpretations of the features in the feature graph. A path in the net represents a consistent succession of characters in the net (figure 12).

In [172], Simon proposes an off-line cursive script recognition scheme. The features are regularities, which are defined as uninformative parts and singularities, which are defined as informative strokes about the characters. Stroke trees are obtained after skeletonization. The goal is to match the trees of singularities.

Although it is computationally expensive, relaxation matching is also a popular method in graphical approaches to CR problem [204], [36], [98].

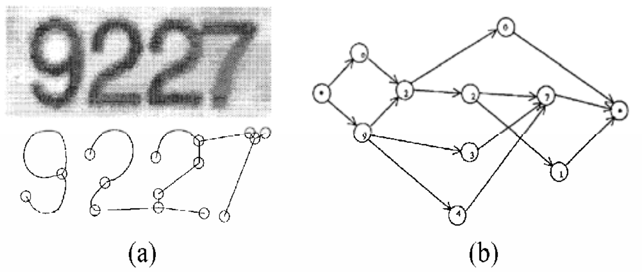


Fig.12. (a) The feature graph of digit string "9227" and (b) its corresponding directed net [163].

4) Neural Networks (NN)

A neural network is defined as a computing architecture that consists of massively parallel interconnection of adaptive 'neural' processors. Because of its parallel nature, it can perform computations at a higher rate compared to the classical techniques. Because of its adaptive nature, it can adapt to changes in the data and learn the characteristics of input signal. A neural network contains many nodes. The output from one node is fed to another one in the network and the final decision depends on the complex interaction of all nodes. In spite of the different underlying principles, it can be shown that most of the neural network architectures are equivalent to statistical pattern recognition methods [161].

Several approaches exist for training of neural networks [79]. These include the error correction, Boltzman, Hebbian and competitive learning. They cover binary and continuous valued input, as well as supervised and unsupervised learning. On the other hand, neural network architectures can be classified into two major groups, namely, feed-forward and feedback (recurrent) networks. The most common neural networks used in the CR systems are the multilayer perceptron of the feed forward networks and the Kohonen's Self Organizing Map (SOM) of the feedback networks.

Multilayer perceptron, proposed by R. Rosenblatt [16] and elaborated by Minsky and Papert [126], is applied in CR by many authors. One example is the feature recognition network proposed by Hussain and Kabuka [75], has two levels detection scheme: The first level is for detection of sub patterns and the second level is for detection of characters. Mohiuddin et al. use multi-network system in hand-printed character recognition by combining contour direction and bending point features [124], which is fed to various neural networks. Neocognitron of Fukushima et al. [48] is a hierarchical network consisting of several layers of alternating neuron-like cells. S-Cells are for feature extracting and C-Cells allow for positional errors in the features. Last layer is the recognition layer. Some of the connections are variable and can be modified by learning. Each layer of S and C cells are called cell planes. This study proposes some techniques for selecting training patterns useful for deformation-invariant recognition of a large number of characters. The feed forward neural network approaches to machine-printed character recognition problem is proved to be successful, in [10], where the neural network is trained with a database of 94 characters and tested in 300 000 characters generated by post script laser printer, with 12 common fonts in varying size. No errors were detected. In this study, Garland et al. propose a two-layer neural network, trained by centroid dithering process. The modular neural network architecture is used for unconstrained handwritten numeral recognition in [143]. The whole classifier is composed of subnetworks. A subnetwork, which contains three layers, is responsible for a class among 10 classes. Another study [167] uses recurrent neural network in order to estimate probabilities for the characters represented in the skeleton of word image (see figure 13). A recent study, proposed by Maragos and Pessoa, incorporates the properties of multilayer perceptron and morphological rank neural networks for handwritten character recognition. They claim that this unified approach gives higher recognition rates than multi-layer perceptron with smaller processing time [150].

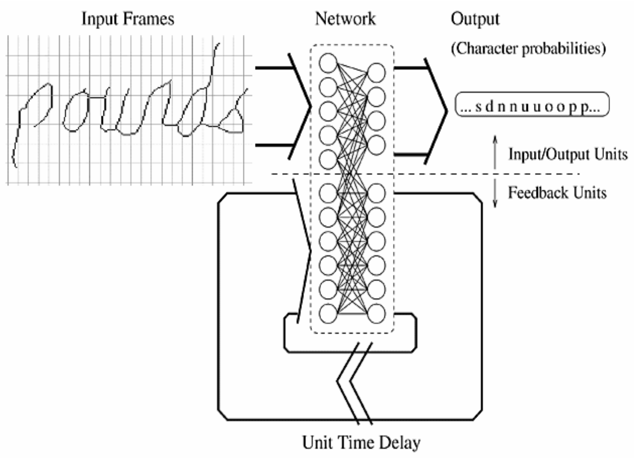


Fig.13. Recurrent neural network in [167].

Most of the recent developments on handwritten CR research are concentrated on Kohonen's Self Organizing Map [95]. SOM integrates the feature extraction and recognition steps in large training set of characters. It can be shown that it is analogous to k-means clustering algorithm. An example of SOM on CR systems is the study of Liou and Yang [115], which presents a self-organization matching approach to accomplish the recognition of handwritten characters, drawn with thick strokes. In [159], Reddy and Nagabhushan propose a combination of modified SOM and Learning Vector Quantization to define three dimensional neural network model for handwritten numeral recognition. They report higher recognition rates with shorter training time than other SOMs reported in the literature. Jabri et al. realized the adaptive-subspace self-organizing map to build a modular classification system for handwritten digit recognition [213].

5) Combined CR Techniques

The above review indicates that, there are many training and recognition methods available for the CR systems. All of the methods have their own superiorities and weaknesses. Now, the question of “May these methods be combined in a meaningful way to improve the recognition results?” appears. In order to answer this question various strategies are developed by combining the CR techniques. Hundreds of these studies can be classified either according to the algorithmic point of view or representational point of view or according to the architecture they use [3]. In this study we suffice to classify the combination approaches in terms of their architecture, in three classes:

1. Serial Architecture,
2. Parallel Architecture,
3. Hybrid Architecture.

Serial architecture feeds the output of a classifier into the next classifier. There are four basic methodologies used in the serial architecture, namely, sequential, selective, boosting and cascade methodologies.

In the sequential methodology, the goal of each stage is to reduce the number of classes in which the unknown pattern may belong. Initially the unknown pattern belongs to one of the total number of classes. The number of probable classes reduces at each stage, yielding the label of the pattern in the final stage [136]. In the selective methodology, initial classifier assigns the unknown pattern into a group of characters that are similar to each other. These groups are further classified in later stages, in a tree hierarchy. At each level of the tree the children of the same parent are similar with respect to a similarity measure. Therefore, the classifiers work from coarse to fine recognition methods in small groups [56]. In the boosting method each classifier handles the classes, which cannot be handled by the previous classifiers [44]. Finally, in the cascade method, the classifiers are connected from simpler to more complex ones. The patterns, which do not satisfy a certain confidence level, are passed to a costlier classifier in terms of features and/or recognition scheme [147].

Parallel architectures combine the result of more than one independent algorithm by using various methodologies. Among many methodologies, the voting [102], Bayesian [84] Dempster-Shafer [205], behavior-knowledge space [74], mixture of experts [76] and stacked generalization [202] are the most representative.

Voting is a democracy-behavior approach based on “the opinion of the majority wins” [102]. It treats classifiers equally without considering their differences in performance. Each classifier represents one score that is either as a whole assigned to one class label or divided into several labels. The label, which receives more than half of the total scores, is taken as the final result. While voting methods are only based on the label without considering the error of each classifier, Bayesian and Dempster-Shafer approaches take these errors into consideration.

The Bayesian approach uses the Bayesian formula to integrate classifiers' decisions. Usually, it requires an independence assumption in order to tackle the computation of the joint probability [84]. The Dempster-Shafer method deals with uncertainty management and incomplete reasoning. It aggregates committed, uncommitted and ignorant beliefs [205]. It allows one to attribute belief to subsets, as well as to individual elements of the hypothesis set. Bayesian and Dempster-Shafer approaches utilize the probabilistic reasoning for classification.

The behavior-knowledge space method has been developed in order to avoid the independence assumption of the individual classifiers [74]. In order to avoid this assumption, the information should be derived from a knowledge space, which can concurrently record the decisions of all classifiers on each learned sample. The knowledge space records the behavior of all classifiers and the method derives the final decisions from the behavior knowledge space.

The mixture of experts' method is similar to the voting method, where decision is weighted according to the input. The experts partition the input space and a gating system decides on the weights for a given input. Thus, the experts are allowed to specialize on local regions of the input space. This is dictated through the used a cost function, based on the likelihood of a normal mixture [76].

Stacked generalization is another extension of the voting method [202], where the outputs of the classifiers are not linearly combined. Instead a combiner system is trained for the final decision.

Lastly, the hybrid architecture is a crossbreeding between the serial and parallel architectures. The main idea is to combine the power of both architectures and to prevent the inconveniences [56].

Examples of the combined approaches may be given as follows:

In [185], a sequential approach based on multifeature and multilevel classification is developed for handwritten Chinese characters. Ten classes of features, such as peripheral shape features, stroke density features and stroke direction features are used in this system. First, a group of classifiers breaks down all the characters into a smaller number of groups; hence the number of candidates for the process in the next step drops sharply. Then, the multilevel character classification method, which is composed of five levels, is employed for the final decision. In the first level, a Gaussian distribution selector is used to select a smaller number of candidates from several groups. From the second level to the fifth one, matching approaches using different features are performed, respectively.

Another example is given by Shridhar and Kimura, where two algorithms are combined for the recognition of unconstrained isolated handwritten numerals [171]. In the first algorithm, a statistical classifier is developed by optimizing a modified quadratic discriminant function derived from the Bayes rule. The second algorithm implements a structural classifier: A tree structure is used to express each number. Recognition is done in a two-pass procedure with different thresholds. The recognition is made either using parallel or serial decision strategies. In a parallel decision strategy, a character is accepted if both passes accept the same result. Otherwise, the results of both algorithms are analyzed for final decision. In a serial decision strategy, if the first algorithm is successful, then the character is accepted. Otherwise, the second algorithm is applied. If it is successful, then the character recognized by the second algorithm is accepted. Otherwise, the final decision is made based on the multiple membership table derived from the results of both algorithms.

In [205], Xu et al. studied the methods of combining multiple classifiers and their application to handwritten recognition. They proposed a serial combination of structural classification and relaxation matching algorithm for the recognition of handwritten zip codes. It is reported that the algorithm has very low error rate and high computational cost.

In [19] Srihari et al. propose a parallel architecture for off-line cursive script word recognition, where they combine three algorithms; template matching, mixed statistical-structural classifier and structural classifier. The results derived from three algorithms are combined in a logical way. Significant increase in the recognition rate is reported.

In [74], Suen et al. derive the best final decision by using behavior-knowledge space method for the recognition of unconstrained handwritten numerals. They use three different classifiers. Their experiments show that the method achieves very promising performances and outperforms voting, Bayesian and Dempster-Shafer approaches.

As an example of the stacked generalization method, a study is proposed in [156] for on-line character recognition. This method combines two classifiers with a neural network. The training ability of neural network works well in managing the conflicts appearing between the classifiers. Here, weighting is subtler as it is applied differently on each class and not equally on all the classifiers score.

In [214], a new kind of neural network -Quantum Neural Network (QNN) is proposed and tested on the recognition of handwritten numerals. QNN combines the advantages of neural modeling and fuzzy theoretic approach. An effective decision fusion system is proposed with high reliability.

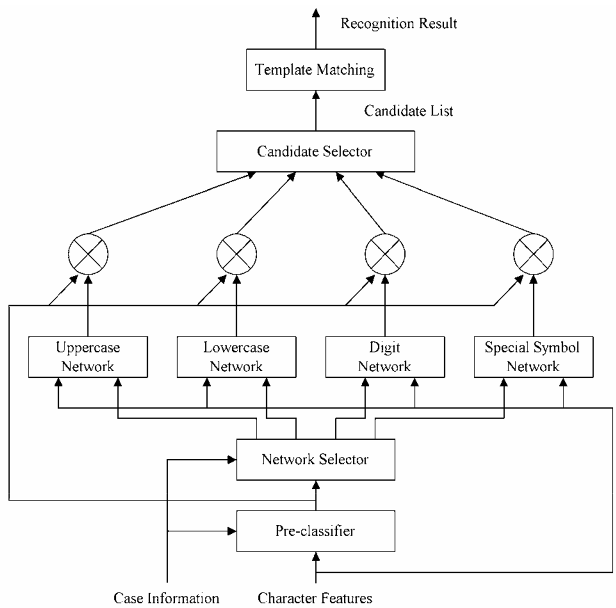


Fig.14. Combined CR System in [56].

A good example of hybrid method is proposed in [56] where IBM research group combines neural network and template matching methods in a complete character recognition scheme (see figure 14). First, the two-stage multi-network (TSMN) classifier identifies the top three candidates. TSMN consists of a bank of specialized networks, each of which is designed to recognize a subset of the entire character set. A pre classifier and a network selector are employed for selectively invoking the necessary specialized networks. Next, the template matching (TM) classifier is invoked to match the input pattern with only those templates in the three categories selected by the TSMN classifier. Template matching distances are used to reorder choices only if the TSMN is not sure about its decision.

E. Post Processing

Until this point no semantic information is considered during the stages of CR. It is well known that humans read by context up to 60% for careless handwriting. While pre-processing tries to "clean" the document in a certain sense, it may remove important information, since the context information is not available at this stage. The lack of context information during the segmentation stage may cause even more severe and irreversible errors, since it yields meaningless segmentation boundaries. It is clear that if the semantic information were available to a certain extent, it would contribute a lot to the accuracy of the CR stages. On the other hand, the entire CR problem is for determining the context of the document image. Therefore utilization of the context information in the CR problem creates a chicken and egg problem. The review of the recent CR research indicates minor improvements, when only shape recognition of the character is considered. Therefore, the incorporation of context and shape information in all the stages of CR systems is necessary for meaningful improvements in recognition rates. This is done in the post processing stage with a feedback to the early stages of CR.

The simplest way of incorporating the context information is the utilization of a dictionary for correcting the minor mistakes of the CR systems. The basic idea is to spell check the CR output and provide some alternatives for the outputs of the recognizer that do not take place in the dictionary [17]. Spelling checkers are available in some languages, like English, German and French etc. String matching algorithms can be used to rank the lexicon words using a distance metric that represents various edition costs [98]. Statistical information derived from the training data and the syntactic knowledge such as N-grams improves the performance of the matching process [64], [87]. In some applications, the context information confirms the recognition results of the different parts in the document image. In automatic reading of bank checks, the inconsistencies between the legal and the courtesy amount can be detected and the recognition errors can be potentially corrected [86].

However, the contextual post-processing suffers from the drawback of making unrecoverable OCR decisions. In addition to the use of dictionary, a well-developed lexicon and a set of orthographic rules contribute a great deal to the recognition rates, in word and sentence levels. In word level, lexicon-driven matching approaches avoid making unrecoverable decisions at the post-processing stage by bringing the context information earlier in the segmentation and recognition stages. A lexicon of words with a knowledge base is used during or after the recognition stage for verification and improvement purpose. A common technique for lexicon driven recognition is to represent the word image by a segmentation graph and match each entry in the lexicon against this graph [90], [33], [9]. In this approach, the dynamic programming technique is often used to rank every word in the lexicon. The word with the highest rank is chosen as the recognition hypothesis. In order to speed up the search process, the lexicon can be represented by a trie data structure or by hash tables [32], [40]. Figure 15 shows an example of segmentation result, corresponding segmentation graph and the representation of the words in the lexicon by a trie structure.

In sentence level, the resulting sentences obtained from the output of the recognition stage can be further processed through parsing in the post processing stage to increase the recognition rates [60], [177], [91]. The recognition choices produced by a word recognizer can be represented by a graph as in the case of word level recognition. Then, the grammatically correct paths over this graph are determined by using syntactic knowledge. However, post processing in sentence level is rather in its infancy, especially for languages other than English, since it requires extensive research in linguistics and formalism in the field of artificial intelligence.

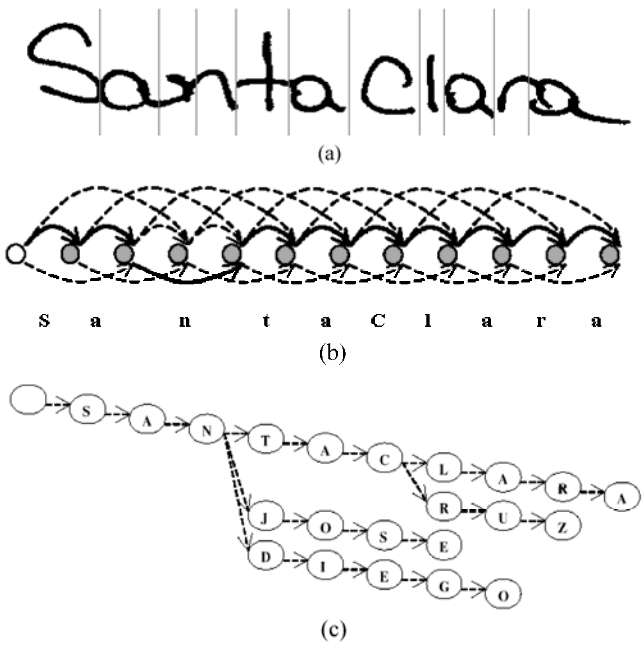


Fig.15. (a) Segmentation result, (b) corresponding segmentation graph and (c) the representation of the lexicon words by a trie structure in [32].

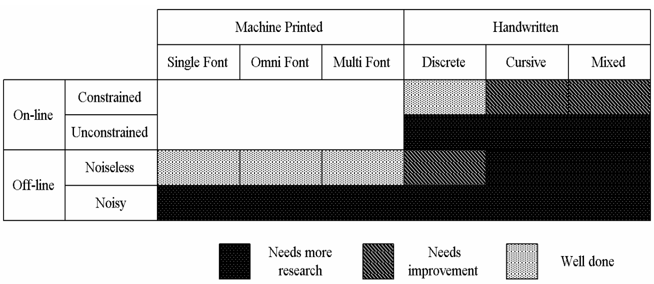
V. Discussion

In this study, we have overviewed the main approaches used in the CR field. Our attempt was to bring out the present status of CR research. Although each of the methods summarized above have their own superiorities and drawbacks, the presented recognition results of different methods seem very successful. Most of the recognition accuracy rates reported is over 85%. However, it is very difficult to make a judgment about the success of the results of recognition methods, especially in terms of recognition rates, because of different databases, constraints and sample spaces. In spite of all the intensive research effort, numerous journal articles, conference proceedings and patents, none of the proposed methods solve the CR problem out of the laboratory environment without putting constraints. The answer to the question "Where are we standing now? " is summarized in Table: 2.

For texts which are handwritten under poor conditions or for free style handwriting, there is still an intensive need in almost all the stages of CR research (solid gray areas). The proposed methods shift towards the hybrid architectures, which utilize, basically, HMM together with the support of structural and statistical techniques. The discrete and clean handwriting on high quality paper or on tablet can be recognized with a rate above 85%, when there is a limited vocabulary (diagonal areas). A popular application area is number digit or limited vocabulary form (bank checks, envelopes and forms designed for specific applications) recognition. These applications may require intensive pre-processing, for skew detection, thinning and base line extraction purposes. Most of the systems use either neural networks or Hidden Markov Models in the recognition stage. Finally, the output of a laser-quality printing device and neat discrete handwriting can be recognized with a rate above 95% (vertical areas). Two leading commercial products, Omni-Page Pro and Recognita, can read complex pages printed on high quality papers, containing the mixture of fonts, without user training.

The best OCR packages in the market use combined techniques based on neural networks for machine- printed characters. Although the input is clean, they require sophisticated pre-processing techniques for noise reduction and normalization. External segmentation is the major task for page layout decomposition. Projection profile is sufficient for character segmentation. Language analysis is used for final confirmation of the recognition stage, providing alternatives in case of a mismatch. Few of the systems, such as Recognita Plus, work on Greek and Cyrillic in addition to Latin alphabet, for limited vocabulary applications. OCR for alphabets other than Latin and for languages other than English, French and German remains mostly in the research arena even for Chinese and Japanese, which have some commercial products.

TABLE II
CURRENT STUDIES IN CR STUDIES



A number of weaknesses, which exist in the proposed systems, can be summarized as follows:

1. In all of the proposed methods, the studies on the stages of CR have come to a point where the improvements are marginal with the current research directions. The stages are mostly based on the shape extracting and recognition techniques and ignore the semantic information. Incorporation of the semantic information is not well explored. In most cases, it is too late to correct all the errors, which propagates through the stages of the CR, in the post processing stage. This situation implies the need of a global change in the approaches for free style handwriting.
2. A major difficulty lies behind the lack of the noise model, over all the stages. Therefore, many assumptions and parameters of the algorithms are set by trial and error, at the initial phase. Unless there are strict constrains about the data and the application domain, the assumptions are not valid even for small changes out of the laboratory environment.
3. Handwriting generation involves semantic, syntactic and lexical information, which is converted into a set of symbols to generate the pen tip trajectory from a predefined alphabet. The available techniques suffer from the lack of characterizing the handwriting generation and the perceptual process in reading, which consists of many complicated phenomena. For example, none of the proposed systems take into account contextual anticipatory phenomena, which lead to co-articulations and the context effects on the writing [153]. The sequence of cursive letters is not produced in serial manner, but parallel articulatory activity occurs. The production of a character is thus affected by the production of the surrounding characters and thus the context.
4. In most of the methods, recognition is isolated from training. The large amount of data is collected and used to train the classifier prior to the classification. Therefore, it is not easy to improve the recognition rate using the knowledge obtained from the analysis of the recognition errors.
5. Selection of the type and the number of features is done by heuristics. It is well known that design of the feature space depends on the training and recognition method. On the other hand, the performance of the recognizer highly depends on the selected features. This problem cannot be solved without evaluating the system with respect to the feature space and the recognition scheme. There exist no evaluation tools for measuring the performance of the stages as well as the overall performance of the system, indicating the source of errors for further improvements.
6. Although color scanners and tablets enable data acquisition with high resolution, there is always a trade-off between the data acquisition quality and complexity of the algorithms, which limits the recognition rate.

Let us evaluate the methods according to the information utilization:

1. Neither the structural nor the statistical information can represent a complex pattern alone. Therefore, one needs to combine statistical and structural information supported by the semantic information.
2. Neural networks or Hidden Markov Models are very successful in combining statistical and structural information for many pattern recognition problems. They are comparatively resistant to distortions. But, they have a discrete nature in the matching process, which may cause drastic mismatching. In other words, they are flexible in training and recognizing samples in classes of reasonably large within-class variances, but they are not continuous in representing the between-class variances. The classes are

separate with their own samples and have models without any relations among them.

3. Template matching methods deal with a character as a whole in the sense that an input plane is matched against a template constrained on X-Y plane. This makes the procedure very simple and the complexity of character shape is irrelevant. But, it suffers from the sensitivity to noise and is not adaptive to variations in writing style. The capabilities of human reasoning are better captured by flexible matching techniques than by direct matching. Some features are tolerant to distortion and take care of style variations, rotation and translation to some extent.

4. The characters are natural entities. They require complex mathematical descriptions to obey the mathematical constraint set of the formal language theory. Imposing a strict mathematical rule on the pattern structure is not particularly practical in character recognition, where intra-class variations are very large.

The followings are some suggestions for future research directions:

1. Five stages of the CR given in this study reached to a plateau, that no more improvement could be achieved by using the current image processing and pattern recognition methodologies. In order to improve the overall performance, higher abstraction levels is to be used for modeling and recognizing the handwriting. This is possible by adopting the artificial intelligence methodologies to the CR field. The stages should be governed by a set of advanced shells utilizing the decision-making, expert systems and knowledge base tools, which incorporate semantic and linguistic information. These tools should provide various levels of abstractions for characters, words and sentences.

2. For a real life CR problem, we need techniques to describe a large number of similar structures of the same category while allowing distinct descriptions among categorically different patterns. It seems that a combined CR model is the only solution to practical character recognition problems.

3. HMM is very suitable for modeling the linguistic information as well as the shape information. The best way of approaching the CR problem is to combine the HMMs at various levels for shape recognition and generating grammars for words and sentences. The HMM classifiers of the hybrid CR system should be gathered by an intelligent tool for maximizing the recognition rate. This may be possible if one achieves to integrate the top-down linguistic models and bottom up character recognition schemes by combining various HMMs, which are related to each other through this tool. The tool may consist of a flexible high-level neural network architecture, which provides continuous description of the classes.

4. Design of the training set should be handled systematically, rather than putting up the available data in the set. Training sets should have considerable size and contain random samples including poorly written ones. Decision of the optimal number of samples from each class, the samples that minimizes within class variance and maximizes between class variance, should be selected according to a set of cost functions. The statistics, relating to the factors such as character shape, slant and stroke variations and the occurrences of characters in a language and interrelations of characters, should be gathered.

5. The feature set should represent the models for handwriting generation and perception, which is reflected to the structural and statistical properties properly so that they are sensitive to intra-class variances and insensitive to inter-class variances.

6. Training and recognition processes should use the same knowledge base. This knowledge base should be built incrementally with several analysis levels by a cooperative effort from the human expert and the computer. Methodologies should be developed to combine knowledge from different sources and to discriminate between correct and incorrect recognition. Evaluation tools should be developed to measure and improve the results of the CR system.

7. The methods on the classical shape-based recognition techniques have reached a plateau for the maximum recognition rate, which is lower than the practical needs. Generally speaking, the future research should focus on the linguistic and contextual information for further improvements.

In conclusion, a lot of efforts are spent in the research on CR. Some major improvements are obtained. Can the machines read human writing with the same fluency as human? Not yet!

ACKNOWLEDGMENT

The authors would like to thank Mr. Kuzeyhan Ozdemir, who initiated the CR studies and contributed a lot to the progress of this paper, during his MS studies. Thanks are also to Mr. Paul Henry Leech for his suggestions and to the referees for their comments.

REFERENCES

[1] I. Abulhaiba, P. Ahmed, "A Fuzzy Graph Theoretic Approach To Recognize The Totally Unconstrained Handwritten Numerals", *Pattern Recognition*, vol.26, no.9, pp.1335-1350, 1993.

[2] O. Akindele, A. Belaid, "Page Segmentation by Segment Tracing", in *Proc. 2nd Int. Conf. Document Analysis and Recognition*, pp.314-344, Japan, 1993.

[3] E. Alpaydin, "Techniques for Combining Multiple Learners", in *Proc. Engineering of Intelligence' 98*, vol.2, pp.6-12, ICSC Press, 1998.

[4] A. Amin, "Off-line Arabic Character Recognition: The State of The Art", *Pattern Recognition*, vol.31, no.5, pp. 517-530, 1998.

[5] C. Arcelli, "A One Pass Two Operation Process To Detect The Skeletal Pixels On The 4-distance Transform", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.11, no.4, pp.411-414, 1989.

[6] N. Arica, F. T. Yarman Vural, "One Dimensional Representation Of Two Dimensional Information For HMM Based Handwritten Recognition", *Pattern Recognition Letters*, vol.21 (6-7), pp.583-592, 2000.

[7] N. Arica, F. T. Yarman-Vural, "A New Scheme for Off-Line Handwritten Connected Digit Recognition", in *Proc. 14th Int. Conf. Pattern Recognition,* pp. 1127-1131, Brisbane, Australia, 1998.

[8] A. Atici, F. Yarman-Vural, "A Heuristic Method for Arabic Character Recognition", *Journal of Signal Processing*, vol.62, pp.87-99, 1997.

[9] E. Augustin, O. Baret, D. Price, S. Knerr, "Legal Amount Recognition on French Bank Checks Using a Neural Network-Hidden Markov Model Hybrid", in *Proc. 6th Int. Workshop Frontiers in Handwriting Recognition*, Korea, 1998.

[10] H. I. Avi-Itzhak, T. A. Diep, H. Gartland, "High Accuracy Optical Character Recognition Using Neural Network with Centroid Dithering", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.17, no.2, pp.218-228, 1995.

[11] H. S. Baird, "Document Image Defect Models", in *Proc. Int. Workshop Syntactical and Structural Pattern Recognition*, pp. 38-47, 1990.

[12] O. Baruch, "Line Thinning By Line Following", *Pattern Recognition Letters*, vol.8, no.4, pp.271-276, 1989.

[13] I. Bazzi, R. Schwartz, J. Makhoul, "An Omnifont Open Vocabulary OCR System for English and Arabic", *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol.21, no.6, pp.495-504, 1999.

[14] A. Belaid, J.P. Haton, "A Syntactic Approach for Handwritten Mathematical Formula Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.6, pp. 105-111, 1984.

[15] S. O. Belkasim, M. Shridhar, M. Ahmadi, "Pattern Recognition with Moment Invariants: A comparative Survey", *Pattern Recognition*, vol.24, no.12, pp.1117-1138, 1991.

[16] H. D. Block, B. W. Knight, F. Rosenblatt, "Analysis of A Four Layer Serious Coupled Perceptron", *II. Rev. Modern Physics*, vol.34, pp.135-152, 1962.

[17] M. Bokser, "Omnifont Technologies", *Proc. of the IEEE,* vol.80, no.7, pp.1066-1078, 1992.

[18] D. Bouchaffra, V. Govindaraju, S. N. Srihari, "Postprocessing of Recognized Strings Using Nonstationary Markovian Models", *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol.21, no.10, pp. 990-999, 1999.

[19] R. M. Bozinovic, S. N. Srihari, "Off-line Cursive Script Word Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.11, no.1, pp.68-83, 1989.

[20] M. K. Brown, S. Ganapathy, "Preprocessing Techniques for Cursive Script Word Recognition", *Pattern Recognition*, vol.16, no.5, pp.447-458, 1983.

[21] H. Bunke, M. Roth, E. G. Schukat-Talamazzani, "Off-line Recognition of Cursive Script Produced by Cooperative Writer", in *Proc. 12th Int. Conf. Pattern Recognition*, pp. 146-151, Jerusalem, Israel, 1994.

[22] C. J. C. Burges, J. I. Be, C. R. Nohl, "Recognition of Handwritten Cursive Postal Words using Neural Networks", in *Proc. USPS 5th Adv. Technology Conf.*, pp.117-129, 1992.

[23] M. Cannon, J. Hockberg, P. Kelly, "Quality Assessment and Restoration of Typewritten Document Images", *Int. Journal Document Analysis and Recognition*, vol.2, no.2/3, pp.80-89, 1999.

[24] P. Carrieres, R. Plamondon, "An Interactive Handwriting Teaching Aid", in *Advances in Handwriting and Drawing : A Multidisciplinary Approach*, C. Faure, G. Lorette, A. Vinter, P. Keuss Eds., Paris, 1994, pp.207-239.

[25] R. G. Casey, E. Lecolinet, "A Survey of Methods and Strategies in Character Segmentation", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.7, pp.690-706, 1996.

[26] R. G. Casey, G. Nagy, "Recursive Segmentation and Classification of Composite Patterns", in *Proc. 6th Int. Conf. Pattern Recognition*, pp.1023-1031, Munchen, Germany,1982.

[27] K. R. Castleman, *Digital Image Processing*, Prentice Hall, 1996.

[28] F. Cesarini, M. Gori, S. Mariani, G. Soda, "Structured Document Segmentation and Representation by The Modified X-Y Tree", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp. 563-566, Bangalore, India, 1999.

[29] M. Y. Chen, A. Kundu, J. Zhou, "Off-line Handwritten Word Recognition Using a Hidden Markov Model Type Stochastic Network", *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol.16, pp.481-496, 1994.

[30] M. Y. Chen, A. Kundu, S. N. Srihari, "Variable Duration Hidden Markov Model and Morphological Segmentation for Handwritten Word Recognition", *IEEE Trans. Image Processing*, vol.4, pp.1675-1688, 1995.

[31] C. Chen, J. DeCurtins, "Word Recognition in a Segmentation-free Approach to OCR", in *Proc. 2nd Int. Conf. Document Analysis and Recognition*, pp.573-579, Japan, 1993.

[32] D. Y. Chen, J. Mao, K. Mohiuddin, "An Efficient Algorithm For Matching A Lexicon With A Segmentation Graph", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.543-546, Bangalore, India, 1999.

[33] W. T. Chen, P. Gader, H. Shi, "Lexicon-Driven Handwritten Word Recognition Using Optimal Linear Combination of Order Statistics", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.21, no.1, pp.77-82, 1999.

[34] M. Chen, X. Ding, "A Robust Skew Detection Algorithm For Grayscale Document Image", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.617-620, Bangalore, India, 1999.

[35] F. H. Cheng, W. H. Hsu, C. A. Chen, "Fuzzy Approach to Solve the Recognition Problem of Handwritten Chinese Characters", *Pattern Recognition*, vol.22, no.2, pp.133-141, 1989.

[36] H. Cheng, W. H. Hsu, M. C. Kuo, "Recognition of Handprinted Chinese Characters via Stroke Relaxation", *Pattern Recognition*, vol. 26, no.4, pp.579-593, 1993.

[37] Y. C. Chim, A. A. Kassim, Y. Ibrahim, "Character Recognition Using Statistical Moments", *Image and Vision Computing*, vol.17, pp.299-307, 1999.

[38] K. Chung, J. Yoon, "Performance Comparison of Several Feature Selection Methods Based on Node Pruning in Handwritten Character Recognition", in *Proc. 4th Int. Conf. Document Analysis and Recognition*, pp.11-15, Ulm, Germany, 1997.

[39] M. Cote, E. Lecolinet, M. Cheriet, C. Y. Suen, "Reading of Cursive Scripts Using A Reading Model and Perceptual Concepts, The PERCEPTO System", *Int. Journal Document Analysis and Recognition*, vol.1, no.1, pp.3-17, 1998.

[40] A. Dengel, R. Hoch, F. Hones, T. Jager, M. Malburg, A. Weigel, "Techniques for Improving OCR Results", *Handbook of Character Recognition and Document Image Analysis*, World Scientific, H. Bunke, P.S.P. Wang Eds. pp.227-254, 1997.

[41] P. A. Devijer, J. Kittler, *Pattern Recognition: A Statistical Approach* , Prentice Hall, 1982.

[42] D. Doermann, "Page Decomposition and Related Research", in *Proc. Symp. Document Image Understanding Technology*, pp. 39-55, Bowie, 1995.

[43] C. Downtown, C.G.Leedham, "Preprocessing and Presorting of Envelope Images for Automatic Sorting Using OCR", *Pattern Recognition*, vol.23, no.3-4, pp.347-362, 1990.

[44] H. Drucker, R. Schapire, P. Simard, "Improving Performance in Neural Networks Using a Boosting Algorithm" in *Advances in NIPS*, S. J. Hanson, J. Cowan, L. Giles, Eds. Morgan Kaufmann, 1993, pp.42-49.

[45] L. D. Earnest, "Machine Reading of Cursive Script", in *Proc. IFIP Congress*, pp.462-466, Amsterdam, 1963.

[46] El-Sheikh, R. M. Guindi, "Computer Recognition of Arabic Cursive Scripts", *Pattern Recognition*, vol.21, no.4, pp.293-302, 1988.

[47] H. Fujisawa, Y. Nakano, K. Kurino, "Segmentation Methods For Character Recognition: From Segmentation To Document Structure Analysis", *Proc. of IEEE*, vol.80, no.7, pp.1079-1092, 1992.

[48] [48] K. Fukushima, N. Wake, "Handwritten Alphanumeric Character Recognition by Neocognition", *IEEE Trans. Neural Networks*, vol. 2, no. 3, pp.355-365, 1991.

[49] [49] K. Fukushima, T. Imagawa, "Recognition and Segmentation of Connected Characters with Selective Attention", *Neural Networks*, vol.6, no.1, pp.33-41, 1993.

[50] P. D. Gader, B. Forester, M. Ganzberger, A. Gillies, B. Mitchell, M. Whalen, and T. Yocum, "Recognition of Handwritten Digits Using Template and Model Matching", *Pattern Recognition*, vol.24, no.5, pp.421-431, 1991.

[51] P. D. Gader, M. A. Khabou, "Automatic Feature Generation for Handwritten Digit Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.12, pp.1256-1262, 1996.

[52] P. D. Gader, M. A. Mohamed, J. M. Keller, "Dynamic Programming Based Handwritten Word Recognition Using the Choquet Fuzzy Integral as the Match Function}, *Journal Electronic Imaging*, vol.5, no.1, pp.15-24, 1996.

[53] M. D. Garris, J.L. Blue, G.T. Candela, D.L. Dimmick, J. Gelst, P.J. Grothers, S.A. Janet, C. L. Wilson, "NIST form Based Handprinted Recognition System", National Institute of Standard and Technology, Geithersburg, MD, Technical Report NISTIR 546920899, July 1994.

[54] A. Gillies, "Cursive Word Recognition Using Hidden Markov Models", in *Proc. U.S. Postal Service Advance Technology Conf.* Washington, D.C. pp.557-563, 1992.

[55] E. Gilloux, "Hidden Markov Models in Handwriting Recognition", *Fundamentals in Handwriting Recognition*, S. Impedovo, Ed. NATO ASI Series F, vol.124, Springer Verlag, 1994.

[56] S. Gopisetty, R. Lorie, J. Mao, M. Mohiuddin, A. Sorin, E. Yair, "Automated forms-processing Software and Services", *IBM Journal of Research and Development*, vol. 40, no. 2, pp.211-230, 1996.

[57] V. K. Govindan, A. P. Shivaprasad, "A Pattern Adaptive Thinning Algorithm, *Pattern Recognition*, vol.20, no.6, pp.623-637, 1987.

[58] V. K. Govindan, A. P. Shivaprasad, "Character recognition- A review", *Pattern Recognition* vol.23, no.7, pp.671-683, 1990.

[59] W. Guerfaii, R. Plamondon, "Normalizing and Restoring On-line Handwriting", *Pattern Recognition*, vol.26, no.3, pp. 418-431, 1993.

[60] D. Guillevic, C. Y. Suen, "Cursive Script Recognition: A Sentence Level Recognition Scheme", in *Proc. Int. Workshop Frontiers in Handwriting Recognition*, pp. 216-223, Taiwan, 1994.

[61] D. Guillevic, C. Y. Suen, "HMM-KNN Word Recognition Engine for Bank Cheque Processing", in *Proc. 14th Int. Conf. Pattern Recognition*, pp. 1526-1529, Brisbane, Australia, 1998.

[62] D. Guillevic, C. Y. Suen, "Recognition of Legal Amounts on Bank Cheques", *Pattern Analysis and Applications*, vol.1, no.1, pp.28-41, 1998.

[63] Z. Guo, R. W. Hall, "Fast Fully Parallel Thinning Algorithms", CVGIP: Image Understanding, vol.55, no.3, pp.317-328, 1992.

[64] I. Guyon, F. Pereira, "Design of a Linguistic Postprocessor Using Variable Memory Length Markov Models", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp.454-457, Montreal, Canada, 1995.

[65] J. Ha, R. Haralick, I. Philips, "Document Page Decomposition by Bounding-Box Projection Technique", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp. 119-122, Montreal, Canada, 1995.

[66] Y. Hamamoto, S. Uchimura, K. Masamizu, S. Tomita, "Recognition of Handprinted Chinese Characters Using Gabor Features", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp.819-823, Montreal, Canada, 1995.

[67] L. D. Harmon, "Automatic Recognition of Print and Scripts", *Proc. of IEEE,* vol.60, no.10, pp.1165-1177, 1972.

[68] A. Hashizume, P. S. Yeh, A. Rosenfeld, "A Method of Detecting the Orientation of Aligned Components", *Pattern Recognition Letters*, vol.4, pp.125-132, 1986.

[69] K. C. Hayes, "Reading Handwritten Words Using Hierarchical Relaxation", *Computer Graphics Image Processing*, vol.14, pp.344-364, 1980.

[70] R. B. Hennis, "The IBM 1975 Optical Page Reader, System Design", *IBM Journal of Research and Development*, pp.346-353, 1968.

[71] [71] O. Hori, "A Video Text Extraction Method for Character Recognition", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.25-29, Bangalore, India, 1999.

[72] J. Hu, S. G. Lim, M.K. Brown, "Writer Independent On-line Handwriting Recognition Using an HMM Approach", *Pattern Recognition*, vol.33, no.1, pp.133-147, 2000.

[73] J. Hu, T. Pavlidis, "A Hierarchical Approach to Efficient Curvilinear Object Searching", *Computer Vision and Image Understanding,* vol.63 (2), pp. 208-220, 1996.

[74] Y. S. Huang, C. Y. Suen, "A Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.17, no.1, pp.90-94, 1995.

[75] B. Hussain, M. Kabuka, "A Novel Feature Recognition Neural Network and its Application to Character Recognition", *IEEE Trans Pattern and Machine Intelligence*, vol. 16, no. 1, pp.99-106, 1994.

[76] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, G. E. Hinton, "Adaptive Mixtures of Local Experts", *Neural Computation*, vol.3, pp.79-87, 1991.

[77] A. K. Jain, Bin Yu, "Document Representation and Its Application to Page decomposition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.20, pp.294-308, 1998.

[78] A. K. Jain, D. Zongker, "Representation and Recognition of Handwritten Digits Using Deformable Templates", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, no.12, pp.1386-1391, 1997.

[79] A. K. Jain, J. Mao, and K.M. Mohiuddin, "Artificial Neural Networks: A Tutorial", *IEEE Computer*, pp.31-44, 1996.

[80] A. K. Jain, K. Karu, "Page Segmentation Using Texture Analysis", *Pattern Recognition*, vol. 29, pp. 743-770, 1996.

[81] A. K. Jain, R.P.W. Duin, J. Mao "Statistical Pattern Recognition: A Review ", IEEE *Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-38, 2000.

[82] B. H. Juang, L. R. Rabiner, "The Segmental K-means Algorithm for Estimating Parameters of Hidden Markov Models", *IEEE Trans. A.S.S.P.*, vol.38, no.9, pp.1639-1641, 1990.

[83] J. Kanai, A. D. Bagdanov, "Projection Profile Based Skew Estimation Algorithm for JPIG Compressed Images", *Int. Journal Document Analysis and Recognition*, vol.1, no.1, pp.43-51, 1998.

[84] H. J. Kang, S. W. Lee, "Combining Classifiers based on Minimization of a Bayes Error Rates", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.398-401, Bangalore, India, 1999.

[85] R. Kashi, J. Hu, W. L. Nelson, W. Turin, "A Hidden Markov Model Approach To On-line Handwritten Signature Verification", *Int. Journal Document Analysis and Recognition*, vol.1, no.2, pp.102-109, 1998.

[86] G. Kaufmann, H. Bunke, "Detection and correction of Recognition Errors in Check Reading", *Int. Journal Document Analysis and Recognition*, vol.2, no.4, pp.211-221, 2000.

[87] F. G. Keenan, L.J. Evett, R.J. Whitrow, "A Large Vocabulary Stochastic Syntax Analyser for Handwriting Recognition", in *Proc. 1st Int. Conf. Document Analysis and Recognition*, pp.794-802, France, 1991.

[88] H. Y. Kim, J. H. Kim, "Handwritten Korean Character Recognition Based on Hierarchical Random Graph Modeling", in *Proc. Int. Workshop Frontiers in Handwriting Recognition*, pp. 577-586, Korea, 1998.

[89] W. Y. Kim, P. Yuan, "A practical Pattern Recognition System for Translation, Scale and Rotation Invariance", in *Proc. Int. Conf. Computer Vision Pattern Recognition*, Seattle, Washington, June 1994.

[90] G. Kim, V. Govindaraju, "A Lexicon Driven Approach to Handwritten Word Recognition for Real-Time Applications", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, no.4, pp.366-379, 1997.

[91] G. Kim, V. Govindaraju, S. Srihari, "Architecture for Handwritten Text Recognition Systems", in *Proc. Int. Workshop Frontiers in Handwriting Recognition*, pp. 113-122, Korea, 1998.

[92] K. Kise, O. Yanagida, S. Takamatsu, "Page Segmentation based on Thinning of Background", in *Proc 13th Int. Conf. Pattern Recognition*, pp. 788-792, Vienna, 1996.

[93] J. Kitler, "Feature Selection and Extraction", *Handbook of Pattern Recognition and Image Processing* T. Young & K. Fu Eds. Academic Press pp. 59-83, 1986.

[94] S. Knerr, V. Anisimov, O. Baret, N. Gorski, D. Price, J. C. Simon, "The A2iA Intercheque System: Courtesy Amount and Legal Amount Recognition for French Checks", *Machine Perception and Artificial Intelligence*, S. Impedova, P.S.P. Wangm H. Bunke, Eds., vol.28, pp.43-86, World Scientific, 1997.

[95] T. Kohonen, "Self Organizing Maps", *Springer Series in Information Sciences*, vol.30, Berlin, 1995.

[96] A. Kornai, K. M. Mohiuddin, S. D. Connell, "An HMM-Based Legal Amount Field OCR System For Checks", *IEEE Trans, Systems, Man and Cybernetics*, pp. 2800-2805, 1995.

[97] A. Kornai, K. M. Mohiuddin, S. D. Connell, "Recognition of Cursive Writing on Personal Checks", in *Proc. Int. Workshop Frontiers in Handwriting Recognition*, pp. 373-378, Essex, 1996.

[98] K. Kukich, "Techniques for Automatically Correcting Words in Text" ACM Computing Surveys, vol.24, no.4, pp.377-439, 1992.

[99] T. Kunango, R. Haralick, I.Phillips, "Nonlinear Local and Global Document Degradation Models", *Int. Journal Imaging Systems and Technologies*, vol.5, no. 4, pp.274-282, 1994.

[100] A. Kundu, Y. He, M.Y. Chen, "Alternatives to Variable Duration HMM in Handwriting Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.20, no.11, pp. 1275-1280, 1998.

[101] A. Kundu, Y. He, "On optimal Order in Modeling Sequence Of Letters in Words Of Common Language As a Markov Chain", *Pattern Recognition*, vol.24, no.7, pp.603 - 608, 1991.

[102] L. Lam C. Y. Suen, "Increasing Experts for Majority Vote in OCR : Theoretical Considerations and Strategies", in *Proc. Int. Workshop Frontiers in Handwriting Recognition*, pp. 245-254, Taiwan, 1994.

[103] L. Lam, C. Y. Suen, "An Evaluation of Parallel Thinning Algorithms, for Character Recognition*", IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 17, no. 9, 1995.

[104] L. Lam, S. W. Lee, C. Y. Suen, "Thinning Methodologies- A Comprehensive Survey", *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol.14, pp.869-885, 1992.
[105] V. F. Leavers, "Which Hough Transform?" *CVGIP : Image Understanding*, vol.58 (2), pp. 250-264, 1993.
[106] E. Lecolinet, "A New Model for Context Driven Word Recognition", in *Proc. SDAIR*, pp.135-139, 1993.
[107] E. Lecolinet, J-P. Crettez, "A Grapheme Based Segmentation Technique for Cursive Script Recognition", in *Proc. 1st Int. Conf. Document Analysis and Recognition*, pp.740-744, Saint-Malo, France, 1991.
[108] S. W. Lee, D.J. Lee, H. S. Park, "A New Methodology for Gray Scale Character Segmentation and Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.10, pp.1045-1050, 1996.
[109] W. L. Lee, K-C. Fan, "Document Image Preprocessing Based on Optimal Boolean Filters", *Signal Processing*, vol.80, no.1, pp.45-55, 2000.
[110] S. W. Lee, Y. J. Kim, "Direct Extraction of Topographic Features for Gray Scale Character Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.17, no.7, pp.724-729, 1995.
[111] S. W. Lee, Y. J. Kim, "Multiresolutional Recognition of Handwritten Numerals with Wavelet Transform and Multilayer Cluster Neural Network", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp.1010-1014, Montreal, Canada, 1995.
[112] R. Legault, C. Y. Suen, "Optimal Local Weighted Averaging Methods in Contour Smoothing", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.7, pp.690-706, 1997.
[113] J. G. Leu, "Edge Sharpening Through Ramp Width Reduction", *Image and Vision Computing*, vol.18, no.6-7, pp.501-514, 2000.
[114] X. Li, W. Oh, J. Hong, W. Gao, "Recognizing Components of Handwritten Characters by Attributed Relational Graphs with Stable Features", in *Proc. 4th Int. Conf. Document Analysis and Recognition*, pp.616-620, 1997.
[115] C. Y. Liou, H.C. Yang, Handprinted Character Recognition Based on Spatial Topology Distance Measurements, ", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.9, pp941-945, 1996.
[116] J. Liu, Y. Tang, Q. He, C. Suen, "Adaptive Document Segmentation and Geometric Relation labeling: Algorithms and Experimental Results", in *Proc 13th Int. Conf. Pattern Recognition*, pp.763-767, Vienna, 1996.
[117] W. Lu, Y. Ren, C. Y. Suen, "Hierarchical Attributed Graph Representation and Recognition of Handwritten Chinese Characters", *Pattern Recognition*, vol. 24, no.7, pp. 617-632, 1991.
[118] S. Madhvanath, E. Kleinberg, V. Govindaraju, S. N. Srihari, "The HOVER System for Rapid Holistic Verification of Off-line Handwritten Phrases", in *Proc. 4th Int. Conf. Document Analysis and Recognition*, pp.855-890, Ulm, Germany,1997.
[119] S. Madhvanath, G. Kim, V. Govindaraju, "Chaincode Contour Processing for Handwritten Word Recognition", *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol.21, no.9, pp.928-932, 1999.
[120] S. Madhvanath, V. Govindaraju, "Reference Lines for Holistic Recognition of Handwritten Words", *Pattern Recognition*, vol.32, no.12, pp. 2021-2028, 1999.
[121] S. A. Mahmoud, I.Abuhabia, R.J.Green, "Skeletonization of Arabic Characters Using Clustering Based Skeletonization Algorithm", *Pattern Recognition*, vol.24, no.5, pp.453-464, 1991.
[122] S. Manke, M. Finke, A. Waibel, "NPen: A Writer Independent, Large Vocabulary On-Line Cursive Handwriting Recognition System", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp.403-408, Montreal, 1995.
[123] J. Mantas, "An Overview of Character Recognition Methodologies", *Pattern Recognition*, vol.19, no.6, pp. 425-430, 1986.
[124] J. Mao, K. M. Mohiuddin, T. Fujisaki, "A Two Stage Multi-Network OCR System with a Soft Pre-Classifier and a Network Selector", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp. 821-828, Montreal, 1995.
[125] A. Meyer, "Pen Computing: A Technology Overview and A Vision", *SIGCHI Bulletin*, vol.27, no.3, pp.46-90, 1995.
[126] M. L. Minsky, S. Papert, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, 1969.
[127] K. T. Miura, R. Sato, S. Mori, "A Method of Extracting Curvature Features and Its Application to Handwritten Character Recognition", in *Proc. 4th Int. Conf. Document Analysis and Recognition,* pp.450-454, Ulm, Germany, 1997.
[128] S. Mo, V. J. Mathews, "Adaptive, Quadratic Preprocessing of Document Images for Binarization", *IEEE Trans. Image Processing*, vol.7, no.7, pp. 992-999, 1998.
[129] M. A. Mohamed, P. Gader, "Generalized Hidden Markov Models - Part II: Application to Handwritten Word Recognition", *IEEE Trans. Fuzzy Systems*, vol.8, no.1, pp.82-95, 2000.
[130] M. A. Mohamed, P. Gader, "Handwritten Word Recognition Using Segmentation-Free Hidden Markov Modeling and Segmentation Based Dynamic Programming Techniques", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.5, pp.548-554, 1996.
[131] K. M. Mohiuddin, J. Mao, "A Comparative Study Of Different Classifiers For Handprinted Character Recognition", *Pattern Recognition in Practice IV*, pp. 437-448, 1994.
[132] S. Mori, C. Y. Suen, K. Yamamoto, "Historical Review of OCR Research and Development", *IEEE-Proceedings*, vol.80, no.7, pp.1029-1057, 1992.
[133] S. Mori, H. Nishida, H. Yamada, *Optical Character Recognition*, Wiley, 1999.
[134] K. Mori, I. Masuda, "Advances in recognition of Chinese characters", in *Proc. 5th Int. Conf. Pattern Recognition,* Miami, pp.692-702 1980.
[135] S. Mori, K. Yamamoto, M. Yasuda, "Research on Machine Recognition of Handprinted Characters", *IEEE Trans. Pattern Analysis, Machine Intelligence*, vol.6, no.4, pp.386-404, 1984.
[136] M. Nakagava, T. Oguni, A. Homma, "A coarse classification of on-line handwritten characters" in *Proc. 5th Int. Workshop Frontiers in Handwriting Recognition*, pp. 417-420, Essex, England, 1996.
[137] M. Nakagawa, K. Akiyama, "A Linear-time Elastic Matching For Stroke Number Free Recognition of On-line Handwritten Characters", in *Proc. Int. Workshop Frontiers in Handwriting Recognition*, pp. 48-56, Taiwan, 1994.
[138] W. Niblack, *An Introduction to Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1986.
[139] H. Nishida, "Structural Feature Extraction Using Multiple Bases", *Computer Vision and Image Understanding*, vol.62 no1, pp. 78-89, July 1995.
[140] L. O'Gorman, "The Document Spectrum for Page Layout Analysis", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.15, pp.162-173, 1993.
[141] I. S. Oh, C. Y. Suen, "Distance Features For Neural Network-Based Recognition of Handwritten Characters", *Int. Journal Document Analysis and Recognition*, vol.1, no.2, pp.73-88, 1998.
[142] I. S. Oh, J. S. Lee, C. Y. Suen, "Analysis of class separation and Combination of Class-Dependent Features for Handwriting Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.21, no.10, pp.1089-1094, 1999.
[143] I. S. Oh, J. S. Lee, S. M. Choi, K. C. Hong, "Class-expert Approach to Unconstrained Handwritten Numeral Recognition", in *Proc.5th Int. Workshop Frontiers in Handwriting Recognition*, pp. 95-102, Essex, England, 1996.
[144] M. Okamoto, K. Yamamoto, "On-line Handwriting Character Recognition Method with Directional Features and Direction Change Features", in *Proc. 4th Int. Conf. Document Analysis and Recognition* , pp.926-930, Ulm, Germany, 1997.
[145] E. Oztop, A. Mulayim, V. Atalay, F. T. Yarman-Vural, "Repulsive Attractive Network for Baseline Extraction on Document Images", *Signal Processing*, vol. 74, no.1, 1999.
[146] M. Parizeau, R. Plamondon, "A Fuzzy-Syntactic Approach to Allograph Modeling for Cursive Script Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.17, no.7, pp. 702-712, 1995.
[147] J. Park, V. Govindaraju, S. N. Srihari, "OCR in A Hierarchical Feature Space", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.22, no.4, pp.400-407, 2000.
[148] T. Pavlidis, "Recognition of Printed Text under Realistic Conditions", *Pattern Recognition Letters*, pp. 326, 1993.
[149] T. Pavlidis,, J. Zhou, "Page Segmentation by White Streams", in *Proc. 1st Int. Conf. Document Analysis and Recognition*, pp.945-953, Saint-Malo, France, 1991.
[150] L. F. C. Pessoa, P. Maragos, "Neural Networks with Hybrid Morphological/Rank/Linear Nodes: A Unifying Framework with

Applications to Handwritten Character Recognition", *Pattern Recognition*, vol.33, pp. 945-960, 2000.

[151] I. T. Philips, "How to Extend and Bootstrap and Existing Data Set with Real-life Degraded Images", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.689-693, Bangalore, India, 1999.

[152] R. Plamandon, D. Lopresti, L. R. B. Schomaker, R. Srihari, "On-line Handwriting Recognition", *Encyclopedia of Electrical and Electronics Eng*., J. G. Webster, ed., vol.15, pp.123-146, New York, Wiley, 1999.

[153] R. Plamandon, S. N. Srihari, "On-line and Off-line Handwriting Recognition: A Comprehensive Survey", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.22, no.1, pp.63-85, 2000.

[154] R. Plamondon, Special Issue on Hand-written Processing and Recognition*, Pattern Recognition*, vol. 26, no.3, 1993.

[155] A. Polesel, G. Ramponi, V. Matthews, "Adaptive Unsharp Masking for Contrast Enhancement", in *Proc. Int. Conf. Image Processing*, vol.1, pp.267-271, 1997.

[156] L. Prevost, M. Milgram, "Automatic Allograph Selection and Multiple Expert Classification for Totally Unconstrained Handwritten Character Recognition", in *Proc. 14th Int. Conf. Pattern Recognition*, pp.381-383, Brisbane, Australia, 1998.

[157] Keith E. Price, "Relaxation Matching Techniques Comparison", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.7, no.5, pp. 617 - 623, 1985.

[158] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proc. of IEEE*, vol.77, pp.257-286, 1989.

[159] N. V. S. Reddy, P. Nagabhushan, "A three Dimensional Neural Network Model for Unconstrained Handwritten Numeral Recognition: A New Approach", *Pattern Recognition,* vol.31, no5, pp511-516, 1998.

[160] J. M. Reinhardt, W. E. Higgins, "Comparison Between The Morphological Skeleton and Morphological Shape Decomposition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.18, no.9, pp.951-957, 1996.

[161] B. Ripley, "Statistical Aspects of Neural Networks", *Networks on Chaos: Statistical and Probabilistic Aspects* ,U. Bornndorff-Nielsen, J. Jensen, W. Kendal, Eds,. Chapman and Hall, 1993.

[162] J. Rocha, T. Pavlidis, "A Shape Analysis Model"*, IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.16, no.4, pp.394-404, 1994.

[163] J. Rocha, T. Pavlidis, "Character Recognition without Segmentation", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.17, no.9, pp. 903-909, 1995.

[164] H. E. S. Said, T. N. Tan, K. D. Baker, "Personal Identification Based on Handwriting", *Pattern Recognition*, vol.33, no.1, pp.149-160, 2000.

[165] J. Saula, M. Pietikainen, "Adaptive Document Image Binarization", *Pattern Recognition*, vol.33, no.2, pp.225-236, 2000.

[166] M. Sekita, K. Toraichi, R. Mori, K. Yamamoto, H. Yamada, "Feature Extraction of Handwritten Japanese Characters by Spline Functions for Relaxation Matching", *Pattern Recognition*, vol.21, no.1, pp. 9-17, 1988.

[167] A. W. Senior, A. J. Robinson, "An Off-Line Cursive Handwriting Recognition", *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol.20, no.3, pp. 309-322, 1998.

[168] J. Serra, "Morphological Filtering: An Overview", *Signal Processing*, vol. 38, no.1, pp.3-11, 1994.

[169] T. Shioyama, H. Y. Wu, T. Nojima, "Recognition Algorithm Based On Wavelet Transform For Handprinted Chinese Characters", in *Proc. 14th Int. Conf. Pattern Recognition*, vol.1, pp.229-232, 1998.

[170] M. Shridhar, A. Badreldin, "High Accuracy Syntactic Recognition Algorithm for Handwritten Numerals", *IEEE Trans. Systems Man and Cybernetics*, vol.15, no.1, pp.152 - 158, 1985.

[171] M. Shridhar, F.Kimura, "Handwritten Numerical Recognition Based On Multiple Algorithms", *Pattern Recognition*, vol.24, no. 10, pp:969-983, 1991.

[172] J. C. Simon, "Off-line Cursive Word Recognition", *Proc. of IEEE*, pp.1150, 1992.

[173] R. M. K. Sinha, B. Prasada, G. F. Houle, M. Babourin, "Hybrid Contextual Text Recognition With String Matching" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.15, no.9, pp.915-925, 1993.

[174] S. Smith, M. Borgoin, K. Sims, H. Voorhees, "Handwritten Character Classification Using Nearest Neighbor in Large Databases", *IEEE Trans. Pattern Recognition and Machine Intelligence,* vol.16, no.9, pp. 915-919, 1994.

[175] Y. Solihin, C.G. Leedham, "Integral Ratio: A New Class of Global Thresholding Techniques for Handwriting Images", *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol.21, no.8, pp.761-768, 1999.

[176] M. Sonka, V.Hlavac, R. Boyle, *Image Processing, Analysis and Machine Vision*, 2nd Ed. PWS Publishing, Brooks/Cole Pub. Company, 1999.

[177] R. K. Srihari, C. M. Baltus, "Incorporating Syntactic Constraints in Recognizing Handwritten Sentences", in *Proc. Int. Joint Conf. Artificial Intelligence*, pp. 1262-1266, France, 1993.

[178] S. N. Srihari, E. J. Keubert, "Integration of Handwritten Address Interpretation Technology into the United States Postal Service Remote Computer Reader System", in *Proc. 4th Int. Conf. Document Analysis and Recognition* , pp.892-896, Ulm, Germany, 1997.

[179] T. Steinherz, E. Rivlin, N. Intrator, "Off-line Cursive script Word Recognition - A survey", *Int. Journal Document Analysis and Recognition*, vol.2, no.2, pp.90-110, 1999.

[180] C. Y. Suen, C. C. Tappert, T. Wakahara, "The State of the Art in On-Line Handwriting Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 8, pp.787-808, 1990.

[181] C. Y. Suen, M. Berthod, S. Mori, "Automatic Recognition of Handprinted Characters - The State of the Art", *Proc. of the IEEE*, vol. 68, no. 4, pp. 469 - 487, 1980.

[182] C. Y. Suen, Tuan A. Mai, "A Generalized Knowledge-Based System for the Recognition of Unconstrained Handwritten Numerals", *IEEE Trans. Systems., Man and Cybernetics*, vol.20, no.4, pp.835-848, 1990.

[183] H. Takahashi, "A Neural Net OCR Using Geometrical And Zonal-Pattern Features", in *Proc. 1st Int. Conf. Document Analysis and Recognition*, 1991.

[184] Y. Tang, H. Ma, X. Mao, D. Liu, C. Suen, "A New Approach to Document Analysis Based on Modified Fractal Signature", in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, pp.567-570, Montreal, Canada, 1995.

[185] Y. Tang, L. T. Tu, J. Liu, S. W. Lee, W. W. Lin, I. S. Shyu, "Off-line Recognition of Chinese Handwriting by Multifeature and Multilevel Classification", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.20, no.5, pp.556-561, 1998.

[186] Y. Tao, Y. Y. Tang, "The Feature Extraction of Chinese Character Based on Contour Information", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.637-640, Bangalore, India, 1999.

[187] C. C. Tappert, "Adaptive on-line handwriting recognition", in *Proc. 7th Int. Conf. Pattern Recognition*, Montreal, Canada, pp. 1004-1007, 1984).

[188] C. C. Tappert, "Cursive Script Recognition by Elastic Matching", *IBM Journal of Research and Development*, vol.26, no.6, pp.765-771, 1982.

[189] M. Tayli, A I. Ai-Salamah, "Building Bilingual Microcomputer System" *Communications of the ACM*, vol.33, no.5, pp.495-504, 1990.

[190] Q. Tian, P. Zhang, T. Alexander, Y. Kim. "Survey: Omnifont Printed Character Recognition", *Visual Communication and Image Processing 91: Image Processing*, pp. 260 - 268, 1991.

[191] ∅. D. Trier, A. K. Jain, "Goal Directed Evaluation of Binarization Methods", *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol.17, pp.1191-1201, 1995.

[192] ∅. D. Trier, A. K. Jain, T. Taxt, "Feature Extraction Method for Character Recognition - A Survey", *Pattern Recognition*, vol.29, no.4, pp.641-662, 1996.

[193] W. H. Tsai, K.S.Fu, "Attributed Grammar- A Tool for Combining Syntactic and Statistical Approaches to Pattern Recognition", *IEEE Trans. System Man and Cybernetics*, vol.10, no.12, pp. 873-885, 1980.

[194] S. Tsujimoto, H. Asada, "Major Components of a Complete Text Reading System", *Proc. of IEEE*, vol.80, no.7 pp.1133, 1992.

[195] D. Tubbs, "A Note on Binary Template Matching", *Pattern Recognition*, vol.22, no.4, pp.359 - 365, 1989.

[196] J. R. Ullmann, "Advances in character recognition", *Applications of Pattern Recognition*. K. S. Fu, Ed., CRC Press, Inc., ch.9, 1986.

[197] J. Wang, J. Jean, "Segmentation of Merged Characters by Neural Networks and Shortest Path", *Pattern Recognition*, vol.27, no.5, pp.649, 1994.

[198] L. Wang, J. Mendel, "A Fuzzy Approach to Handwritten Rotation invariant Character Recognition", *Computer Magazine*, vol.3, pp.145-148, 1992.

[199] S. S. Wang, P. C. Chen, W. G. Lin, "Invariant Pattern Recognition by Moment Fourier Descriptor", *Pattern Recognition*, vol.27, pp.1735-1742, 1994.

[200] K. Wang, Y. Y. Tang, C. Y. Suen, "Multi-Layer Projections for the Classification of Similar Chinese Characters", in *Proc. 9th Int. Conf. Pattern Recognition*, pp. 842-844, Italy, 1988.

[201] T. Watanabe, L Qin, N. Sugie, "Structure Recognition Methods for Various Types of Documents", *Machine Vision and Applications*, vol. 6, pp.163-176, 1993.

[202] D. H. Wolperk, "Stacked Generalization", *Neural Networks*, vol.5, pp. 241-259, 1992.

[203] M. Worring, A. W. M. Smeulders, "Content Based Internet Access to Paper Documents", *Int. Journal Document Analysis and Recognition*, vol.1, no.4, pp., 1999.

[204] S. L. Xie, M. Suk, "On Machine Recognition of Hand-printed Chinese Characters by Feature Relaxation", *Pattern Recognition* vol.21, no.1, pp.1-7, 1988.

[205] L. Xu, A. Krzyzak, C.Y. Suen, "Methods of combining Multiple classifiers and their Application to Handwritten Recognition", *IEEE Trans. Systems Man and Cybernetics*, vol 22, no 3, pp418-435, 1992.

[206] K. Yamamoto, "A. Rosenfeld, Recognition of Handprinted Kanji Characters by a Relaxation Method", in *Proc. 6th Int. Conf. Pattern Recognition*, Germany, pp.395-398, 1982.

[207] Y. Yang, H. Yan, "An Adaptive Logical Method for Binarization of Degraded Document Images", *Pattern Recognition*, vol.33, no.5, pp.787-807, 2000.

[208] J. Yang, X. B. Li, "Boundary Detection Using Mathematical Morphology", *Pattern Recognition Letters*, vol.16, no.12, pp.1287-1296, 1995.

[209] B. A. Yanikoglu, P. Sandon, "Recognizing Off-line Cursive Handwriting", in *Proc. Int. Conf. Computer Vision Pattern Recognition*, pp. 397-403, Seattle, 1994.

[210] F. T. Yarman-Vural, A. Atici, "A Segmentation and Feature Extraction Algorithm for Ottoman Cursive Script", in *Proc. 3rd Turkish Symposium on Artificial Intelligence and Neural Networks*, Turkey, 1994.

[211] F. T. Yarman-Vural, E. Ataman, "Noise Histogram and Cluster Validity for Gaussian Mixtured Data", *Pattern Recognition*, vol.20, no.4, pp.385-401, 1987.

[212] B. Yu, A. K. Jain, "A Robust and Fast Skew Detection Algorithm for Generic Documents", *Pattern Recognition*, 1996.

[213] B. Zhang, M. Fu. H. Yan, M.A. Jabri, "Handwritten Digit Recognition by Adaptive -Subspace Self Organizing Map (ASSOM)", *IEEE Trans. Neural Network*, vol. 10, no.4. 1999.

[214] J. Zhou, Q. Gan, A. Krzyzak, C. Suen, "Recognition of handwritten Numerals by Quantum Neural Network with fuzzy Feature", *Int. Journal Document Analysis and Recognition*, vol.2, no.1, pp 30-36, 1999.

[215] X. Zhu, Y. Shi, S. Wang, "A New Algorithm of Connected Character Image Based on Fourier Transform", in *Proc. 5th Int. Conf. Document Analysis and Recognition*, pp.788-791, Bangalore, India, 1999.