

# Handwritten Character Recognition Using Bayesian Decision Theory

Vijiyakumar, Suresh Joseph

Department of computer science, Pondicherry University  
Pondicherry-605014, India.  
vijiya.kumar@gmail.com, sureshjosephk@yahoo.co.in

**Abstract:** Character recognition (CR) can solve more complex problem in handwritten character and make recognition easier. Handwriting character recognition (HCR) has received extensive attention in academic and production fields. The recognition system can be either online or offline. Offline handwritten character recognition is the sub fields of optical character recognition (OCR). The offline handwritten character recognition stages are preprocessing, segmentation, feature extraction and recognition. Our aim is to improve missing character rate of an offline character recognition using Bayesian decision theory.

**Keywords:** Character recognition, Optical character recognition, Off-line Handwriting, Segmentation, Feature extraction, Bayesian decision theory.

## 1. Introduction

The recognition system can be either on-line or off-line. On-line handwriting recognition involves the automatic conversion of text as it is written on a special digitized or PDA, where a sensor picks up the pen-tip movements as well as pen-up/pen-down switching. That kind of data is known as digital ink and can be regarded as a dynamic representation of handwriting. Off-line handwriting recognition involves the automatic conversion of text in an image into letter codes which are usable within computer and text-processing applications. The data obtained by this form is regarded as a static representation of handwriting.

The aim of character recognition is to translate human readable character to machine readable character. Optical character recognition is a process of translation of human readable character to machine readable character in optically scanned and digitized text. Handwritten character recognition (HCR) has received extensive attention in academic and production fields.

Bayesian decision theory is a fundamental statistical approach that quantifies the tradeoffs between various decisions using probabilities and costs that accompany such decision.

They divided the decision process into the following five steps:

1. Identification of the problem.
2. Obtaining necessary information.
3. Production of possible solution.
4. Evaluation of such solution.
5. Selection of a strategy for performance.

They also include a sixth stage implementation of the decision. In the existing approach missing data cannot be

recognition which is useful in recognition historical data. In our approach we are recognition the missing words using Bayesian classifier. It first classifier the missing words to obtain minimize error. It can recover as much error as possible.

## 2. Related Work

The history of CR can be traced as early as 1900, when the Russian scientist Turing attempted to develop an aid for the visually handicapped [1]. The first character recognizers appeared in the middle of the 1940s with the development of digital computers. The early work on the automatic recognition of characters has been concentrated either upon machine-printed text or upon a small set of well-distinguished handwritten 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 [2].

The commercial character recognizers were available in the 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 [3].

Studies up until 1980 suffered from the lack of powerful computer hardware and data acquisition devices. With the explosion of information technology, the previously developed methodologies found a very fertile environment for rapid growth addition to the statistical methods. The CR research was focused basically on 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 [4] and [3] for off-line and on-line cases, respectively.

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 1990s, image processing and pattern recognition techniques were efficiently combined with artificial intelligence (AI) 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 (NNs), hidden Markov models (HMMs), fuzzy set reasoning, and natural language processing. The recent systems for the machine-printed off-line [2] [5] and limited vocabulary, user-dependent on-line handwritten characters [2] [12] 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.

Bayesian decision Theory (BDT), one of the statistical techniques for pattern classification, to identify each of the large number of black-and-white rectangular pixel displays as one of the 26 capital letters in the English alphabet. The character images were based on 20 different fonts and each letter within 20 fonts was randomly distorted to produce a file of 20,000 unique instances [6].

### 3. Existing System

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: 1) the data acquisition process (on-line or off-line) and 2) the text type (machine-printed or handwritten). No matter in which class the problem belongs, in general, there are five major stages Figure 1 in the CR problem:

- 1) Preprocessing
- 2) Segmentation
- 3) Feature Extraction
- 4) Recognition
- 5) Post processing

#### 3.1. Preprocessing

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. Preprocessing aims to produce data that are easy for the CR systems to operate accurately.

The main objectives of preprocessing 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 preprocessing stage.

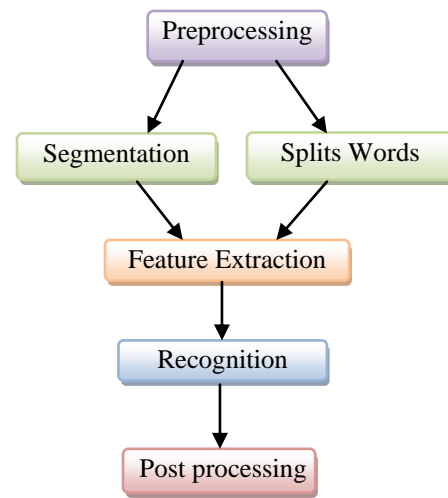


Figure 1. Character recognition

#### 3.1.1 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 CR, it is necessary to eliminate these imperfections. Hundreds of available noise reduction techniques can be categorized in three major groups [7] [8]:

- a) Filtering
- b) Morphological Operations
- c) Noise Modeling

#### 3.1.2 Normalization

Normalization methods aim to remove the variations of the writing and obtain standardized data. The following are the basic methods for normalization [4] [10][16].

- a) Skew Normalization and Baseline Extraction
- b) Slant Normalization
- c) Size Normalization

#### 3.1.3 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 CR requires space domain techniques for preserving the shape information.

a) Threshold: 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 threshold exist: *global* and *local*. Global threshold picks one threshold value for the entire document image which is often based on an estimation of the background level from the intensity histogram of the image. Local (adaptive) threshold use different values for each pixel according to the local area information.

b) 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 1) *pixel wise* and 2) *nonpareil wise* thinning [1]. Pixel wise thinning methods locally and iteratively process the image until one pixel wide skeleton remains. They are very sensitive to noise and may deform the shape of the character. On the other hand, the no 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. In 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. These points are the source of problems. In a nonpareil wise thinning, they are handled with global approaches. A survey of pixel wise and nonpareil wise thinning approaches is available in [9].

### 3.2. Segmentation

The preprocessing stage yields a “clean” document in the sense that a sufficient amount of shape information, high compression, and low noise on a normalized image is obtained. The next stage is segmenting the document into its subcomponents. 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: external segmentation, which is the isolation of various writing units, such as paragraphs, sentences, or words, and internal segmentation, which is the isolation of letters, especially in cursively written words.

1) *External Segmentation*: It is the most critical part of the document analysis, which is a necessary step prior to the off-line CR. 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.), and 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.) [12].

2) *Internal Segmentation*: 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 [13] is Explicit Segmentation, Implicit Segmentation and Mixed Strategies.

### 3.3. Feature Extraction

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 [14]. A good survey on feature extraction methods for CR can be found [15]. In the following, hundreds of document image representations methods are categorized into three major groups are Global Transformation and Series Expansion, Statistical Representation and Geometrical and Topological Representation.

### 3.4. Recognition Techniques

CR systems extensively use the methodologies of pattern recognition, which assigns an unknown sample into a predefined class. Numerous techniques for CR can be investigated in four general approaches of pattern recognition, as suggested in [16] are Template matching, Statistical techniques, and Structural techniques and Neural networks.

### 3.5. 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 preprocessing 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.

## 4. The proposed System Architecture

The proposed research methodology for off-line cursive handwritten characters is described in this section as shown in Figure 2.

### 4.1 Preprocessing

There exist a whole lot of tasks to complete before the actual character recognition operation is commenced. These preceding tasks make certain the scanned document is in a suitable form so as to ensure the input for the subsequent recognition operation is intact. The process of refining the scanned input image includes several steps that include: Binarization, for transforming gray-scale images in to black & white images, scraping noises, Skew Correction-performed to align the input with the coordinate system of the scanner and etc., The preprocessing stage comprise three steps:

- (1) Binarization
- (2) Noise Removal

(3) Skew Correction

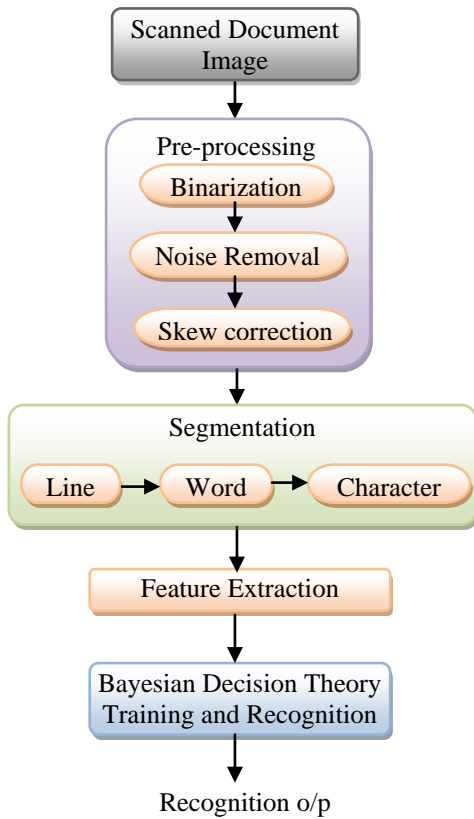


Figure 2. Proposed System Architecture

4.1.1 Binarization

Extraction of foreground (ink) from the background (paper) is called as threshold. Typically two peaks comprise the histogram gray-scale values of a document image: a high peak analogous to the white background and a smaller peak corresponding to the foreground. Fixing the threshold value is determining the one optimal value between the peaks of gray-scale values [1]. Each value of the threshold is tried and the one that maximizes the criterion is chosen from the two classes regarded as the foreground and back ground points.

4.1.2 Noise Removal

The presence of noise can cost the efficiency of the character recognition system; this topic has been dealt extensively in document analysis for typed or machine-printed documents. Noise may be due the poor quality of the document or that accumulated whilst scanning, but whatever is the cause of its presence it should be removed before further Processing. We have used median filtering and Wiener filtering for the removal of the noise from the image.

4.1.3 Skew Correction

Aligning the paper document with the co-ordinate system of the scanner is essential and called as skew correction. There exist a myriad of approaches for skew correction covering correlation, projection, profiles, Hough transform and etc.

For skew angle detection Cumulative Scalar Products (CSP) of windows of text blocks with the Gabor filters at different orientations are calculated. Alignment of the text line is used as an

important feature in estimating the skew angle. We calculate CSP for all possible 50X50 windows on the scanned document image and the median of all the angles obtained gives the skew angle.

4.2 Segmentation

Segmentation is a process of distinguishing lines, words, and even characters of a hand written or machine-printed document, a crucial step as it extracts the meaningful regions for analysis. There exist many sophisticated approaches for segmenting the region of interest. Straight-forward, may be the task of segmenting the lines of text in to words and characters for a machine printed documents in contrast to that of handwritten document, which is quiet difficult. Examining the horizontal histogram profile at a smaller range of skew angles can accomplish it. The details of line, word and character segmentation are discussed as follows.

4.2.1 Line Segmentation

Obviously the ascenders and descenders frequently intersect up and down of the adjacent lines, while the lines of text might itself flutter up and down. Each word of the line resides on the imaginary line that people use to assume while writing and a method has been formulated based on this notion shown fig.3.

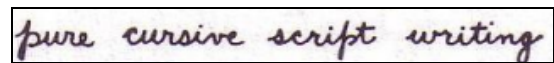


Figure 3. Line Segmentation

The local minima points are calibrated from each Component to approximate this imaginary baseline. To calculate and categorize the minima of all components and to recognize different handwritten lines clustering techniques are deployed.

4.2.2 Word and Character Segmentation

The process of word segmentation succeeds the line separation task. Most of the word segmentation issues usually concentrate on discerning the gaps between the characters to distinguish the words from one another other. This process of discriminating words emerged from the notion that the spaces between words are usually larger than the spaces between the characters in fig 4.

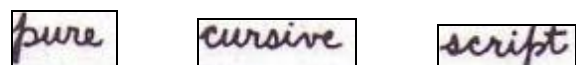


Figure 4. Word Segmentation

There are not many approaches to word segmentation issues dealt in the literature. In spite of all these perceived conceptions, exemptions are quiet common due to flourishes in writing styles with leading and trailing ligatures. Alternative methods not depending on the one-dimensional distance between components, incorporates cues that humans use. Meticulous examination of the variation of spacing between the adjacent characters as a function of the corresponding characters themselves helps reveal the writing style of the author, in terms of spacing. The segmentation scheme comprises the notion of expecting greater spaces between characters with leading and trailing ligatures.

Recognizing the words themselves in textual lines can itself help lead to isolation of words. Segmentation of words in to its constituent characters is touted by most recognition methods. Features like ligatures and concavity are used for determining the segmentation points.

### 4.3 Feature Extraction

The size inevitably limited in practice, it becomes essential to exploit optimal usage of the information stored in the available database for feature extraction. Thanks to the sequence of straight lines, instead of a set of pixels, it is attractive to represent character images in handwritten character recognition. Whilst holding discriminated information to feed the classifier, considerable reduction on the amount of data is achieved through vector representation that stores only two pairs of ordinates replacing information of several pixels. Vectorization process is performed only on basis of bi-dimensional image of a character in off-line character recognition, as the dynamic level of writing is not available. Reducing the thickness of drawing to a single pixel requires thinning of character images first. Character before and after Thinning After streamlining the character to its skeleton, entrusting on an oriented search process of pixels and on a criterion of quality of representation goes on the vectorization process. The oriented search process principally works by searching for new pixels, initially in the same direction and on the current line segment subsequently. The search direction will deviate progressively from the present one when no pixels are traced. The dynamic level of writing is retrieved of course with moderate level of accuracy, and that is object of oriented search. Starting the scanning process from top to bottom and from left to right, the starting point of the first line segment, the first pixel is identified. According to the oriented search principle, specified is the next pixel that is likely to be incorporated in the segment. Horizontal is the default direction of the segment considered for oriented search. Either if the distortion of representation exceeds a critical threshold or if the given number of pixels has been associated with the segment, the conclusion of line segment occurs. Computing the average distance between the line segment and the pixels associated with it will yield the distortion of representation. The sequence of straight lines being represented through ordinates of its two extremities character image representation is streamlined finally. All the ordinates are regularized in accordance to the initial width and height of character image to resolve scale Variance.

### 4.4 Bayesian Decision Theories

The Bayesian decision theory is a system that minimizes the classification error. This theory plays a role of a prior. This is when there is priority information about something that we would like to classify.

It is a fundamental statistical approach that quantifies the tradeoffs between various decisions using probabilities and costs that accompany such decisions. First, we will assume that all probabilities are known. Then, we will study the cases where the probabilistic structure is not completely known. Suppose we know  $P(w_j)$  and  $p(x|w_j)$  for  $j = 1, 2, \dots, n$ . and measure the lightness of a fish as the value  $x$ .

Define  $P(w_j|x)$  as the a posteriori probability (probability of the state of nature being  $w_j$  given the measurement of feature value  $x$ ).

We can use the Bayes formula to convert the prior probability to the posterior probability

$$P(w_j|x) = \frac{p(x|w_j)p(w_j)}{p(x)}$$

$$\text{Where } p(x) = \sum_{j=1}^c p(x|w_j) p(w_j)$$

$P(x|w_j)$  is called the likelihood and  $p(x)$  is called the evidence.

Probability of error for this decision

$$P(\text{error}|x) = \begin{cases} P(w_1|x) & \text{if we decide } w_2 \\ P(w_2|x) & \text{if we decide } w_1 \end{cases}$$

Average probability of error

$$P(\text{error}) = \int_{-\infty}^{\infty} p(\text{error}, x) dx$$

$$P(\text{error}) = \int_{-\infty}^{\infty} p(\text{error}, x) p(x) dx$$

Bayes decision rule minimizes this error because

$$P(\text{error}|x) = \min \{P(w_1|x), P(w_2|x)\}$$

Let  $\{w_1, \dots, w_c\}$  be the finite set of  $c$  states of nature (classes, categories). Let  $\{\alpha_1, \dots, \alpha_a\}$  be the finite set of  $a$  possible actions. Let  $\lambda(\alpha_i|w_j)$  be the loss incurred for taking action  $\alpha_i$  when the state of nature is  $w_j$ . Let  $x$  be the

$D$ -component vector-valued random variable called the feature vector.

$P(x|w_j)$  is the class-conditional probability density function.  $P(w_j)$  is the prior probability that nature is in state  $w_j$ . The posterior probability can be computed as

$$P(w_j|x) = \frac{p(x|w_j)p(w_j)}{p(x)}$$

$$\text{Where } p(x) = \sum_{j=1}^c p(x|w_j) p(w_j)$$

Suppose we observe  $x$  and take action  $\alpha_i$ . If the true state of nature is  $w_j$ , we incur the loss  $\lambda(\alpha_i|w_j)$ .

The expected loss with taking action  $\alpha_i$  is

$$R(\alpha_i|x) = \sum_{j=1}^c \lambda(\alpha_i|w_j) p(w_j|x) \text{ which is also called}$$

the conditional risk.

The general decision rule  $\alpha(x)$  tells us which action to take for observation  $x$ . We want to find the decision rule that minimizes the overall risk

$$R = \int R(\alpha(x)|x) p(x) dx$$

Bayes decision rule minimizes the overall risk by selecting the action  $\alpha_i$  for which  $R(\alpha_i|x)$  is minimum. The resulting minimum overall risk is called the Bayes risk and is the best performance that can be achieved.

### 4.5 Simulations

This section describes the implementation of the mapping and generation model. It is implemented using GUI (Graphical User Interface) components of the Java programming under Eclipse Tool and Database storing data in Microsoft Access.

For given Handwritten image character and convert to Binarization, Noise Remove and Segmentation as shown in Figure 5(a). Then after perform Feature Extraction, Recognition using Bayesian decision theory as shown in Figure5(b).

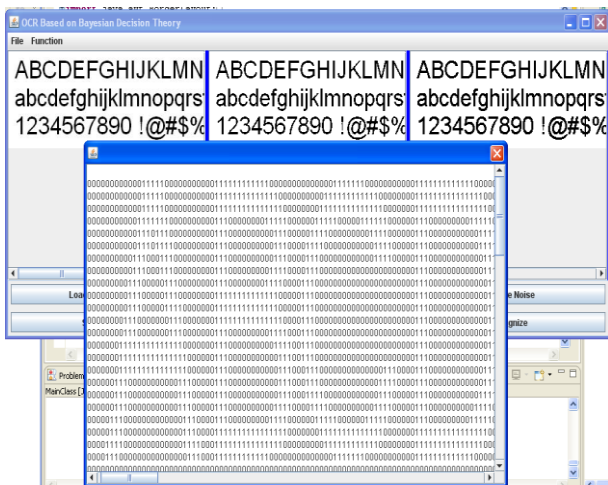


Figure 5(a) Binarization, Noise Remove and Segmentation

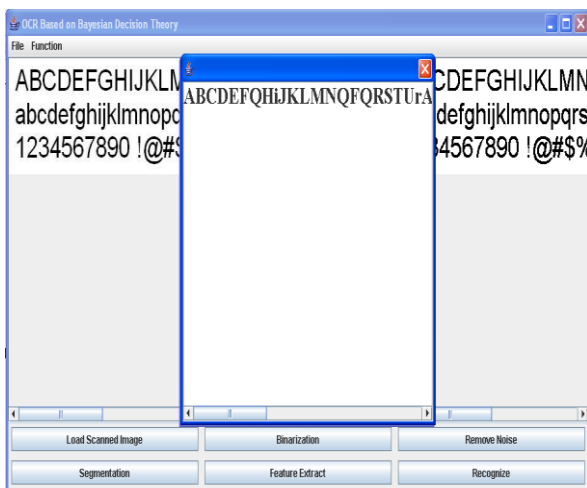


Figure 5(b) Recognition using Bayesian decision theory

## 5. Results and Discussion

This database contains 86,272 word instances from an 11,050 word dictionary written down in 13,040 text lines. We used the sets of the benchmark task with the closed vocabulary IAM-OnDB-t13. There the data is divided into four sets: one set for training; one set for validating the Meta parameters of the training; a second validation set which can be used, for example, for optimizing a language model; and an independent test set. No writer appears in more than one set. Thus, a writer independent recognition task is considered. The size of the vocabulary is about 11K. In our experiments, we did not include a language model. Thus the second validation set has not been used.

Table1. Shows the results of the four individual recognition systems [17]. The word recognition rate is simply measured by dividing the number of correct

recognized words by the number of words in the transcription.

We presented a new Bayesian decision theory for the recognition of handwritten notes written on a whiteboard. We combined two off-line and two online recognition systems. To combine the output sequences of the recognizers, we incrementally aligned the word sequences using a standard string matching algorithm. Evaluation of proposed Bayesian decision theory with existing recognition systems with respect to graph is shown in figure 6.

Table 1. Results of four individuals recognition systems

System	Method	Recognition rate	Accuracy
1st Offline	Hidden Markov Method	66.90%	61.40%
1st Online	ANN	73.40%	65.10%
2nd Online	HMM	73.80%	65.20%
2nd Offline	Bayesian Decision theory	75.20%	66.10%

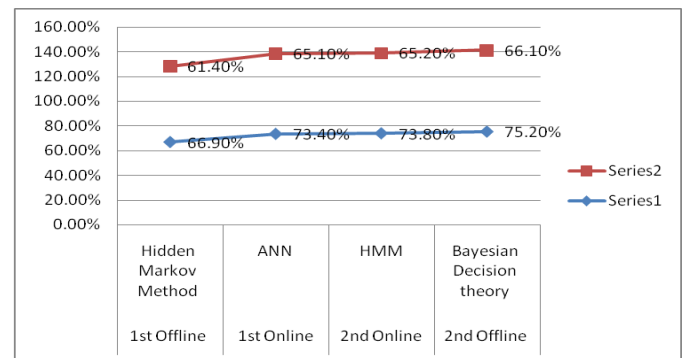


Figure 6 Evaluation of Bayesian decision theory with existing recognition systems

Then each output position the word with the most occurrences has been used as the final result. With the Bayesian decision theory could statistically significantly increase the accuracy.

## 6. Conclusion

We conclude that the proposed approach for offline character recognition, which fits the input character image for the appropriate feature and classifier according to the input image quality. In existing system missing characters can't be identified. Our approach using Bayesian Decision Theories which can classify missing data effectively which decrease error in compare to hidden Markova model. Significantly increases in accuracy levels will found in our method for character recognition

## References

[1] El-Sheikh and R. M. Guindi, "Computer recognition of Arabic cursive scripts," Pattern Recognition., vol. 21, no. 4, pp. 293–302, 1988.

- [2] Alex Graves, Marcus Liwicki, Santiago Fernández, Roman Bertolami, Horst Bunke, and Jürgen Schmidhuber "A Novel Connectionist System for Unconstrained Handwriting Recognition" *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 5, may 2009
- [3] C. Y. Suenf, C. C. Tappert, and T. Wakahara, "The state of the art in on-line handwriting recognition," *IEEE Trans. Pattern Anal. Machine Intel.*, vol. 12, pp. 787–808, Aug. 1990.
- [4] L. Lam, S. W. Lee, and C. Y. Suen, "Thinning methodologies—A comprehensive survey," *IEEE Trans. Pattern Anal. Machine Intel.*, vol. 14, pp. 869–885, Sept. 1992
- [5] R. Plamondon and S. Sridhar. "On-line and Offline Handwriting Recognition: A Comprehensive Survey." *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(1):63–84, 2000.
- [6] Mujtaba Husnain, Shahid Naweed, "English Letter Classification Using Bayesian Decision Theory and Feature Extraction Using Principal Component Analysis" *ISSN 1450-216X Vol.34 No.2*, pp.196-203, 2009.
- [7] R. G. Casey and E. Lecolinet, "A survey of methods and strategies in character segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 690–706, July 2001.
- [8] J. Serra, "Morphological filtering: An overview," *Signal Process.*, vol.38, no. 1, pp. 3–11, 1994.
- [9] C. Downtown and C. G. Leedham, "Preprocessing and presorting of envelope images for automatic sorting using OCR," *Pattern Recognition.*, vol. 23, no. 3–4, pp. 347–362, 1998.
- [10] W. Guerfaii and R. Plamondon, "Normalizing and restoring on-line handwriting," *Pattern Recognition.*, vol. 26, no. 3, pp. 418–431, 1993.
- [11] Ø. D. Trier, A. K. Jain, and T. Text, "Feature extraction method for character recognition—A survey," *Pattern Recognition.*, vol. 29, no. 4, pp. 641–662, 1996.
- [12] R. Munguia, K. Tosca no, G Sanchez, M. Nakano "New Optimized Approach for Written Character Recognition Using Symlest Wavelet" *52nd IEEE International Midwest Symposium on Circuits and Systems 2009*.
- [13] L. O’Gorman, "The document spectrum for page layout analysis," *IEEE Trans. Pattern Anal. Machine Intell.* vol. 15, pp. 162–173, 1993.
- [14] M. Y. Chen, A. Kundu, and J. Zhou, "Off-line handwritten word recognition using a hidden Markov model type stochastic network," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 481–496, May 1994.
- [15] I. S. Oh, J. S. Lee, and C. Y. Suen, "Analysis of class separation and combination of class-dependent features for handwriting recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 1089–1094, Oct. 2002
- [16] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing, Analysis and Machine Vision", 2nd ed. Pacific Grove, CA: Brooks/Cole, 1999.
- [17] Marcus Liwicki, Horst Bunke "Combining On-Line and Off-Line Systems for Handwriting Recognition" *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*.