Stroke-Database Design for Online Handwriting Recognition in Bangla

K. Roy
West Bengal State University, Barasat, Kolkata, India

Abstract: *Handwriting recognition is a difficult task because of the variability involved in the writing styles of different individuals. This gets more complicated in Asian script for its large character set and presence of modified and compound characters. This paper presents a scheme for the online handwriting recognition of one of the major Indic script, Bangla. Online handwriting recognition refers to the problem of interpretation of handwriting input captured as a stream of pen positions using a digitizer or other pen position sensor. Strokes are extracted from the characters. The sequential and dynamical information obtained from the pen movements on the writing pads are used as features in the proposed scheme. These features computed from the strokes, and fed to the MLP classifier for recognition. The characters are identified from its constituent strokes. The system was tested on 21372 Bangla character data and obtained 92.92% accuracy with top 3 choices.*

I. Introduction

Data entry using pen-based devices is gaining popularity in recent times. This is so because machines are getting smaller in size and keyboards are becoming more difficult to use. Also, data entry for Indian scripts having large alphabet size is difficult using keyboard. Moreover, there is an attempt to mimic the pen and paper metaphor by automatic processing of online characters. Work on online character recognition started gaining momentum about forty years ago. Numerous approaches have been proposed in the literature [1] - [9].

Many techniques are available for on-line recognition of English, Arabic, Japanese and Chinese [1] - [9] characters but there are only a few pieces of work [10] – [15], [19] available towards Indian characters although India is a multi-lingual and multi-script country. Connell et al. [11] presented a preliminary study on online Devnagari character recognition. They considered only the main characters that occur in the core-strip neglecting the ascending and the descending parts of the characters. Connell et al. [13] also proposed a work on Devnagari on-line character recognition. Joshi et al. [14] proposed an elastic matching based scheme for on-line recognition of Tamil character recognition. Although there are some work towards on-line recognition of Devnagari and Tamil scripts but the on-line recognition work towards other Indian languages is very few. In this paper we propose a system for the on-line recognition of Bangla handwritten characters. Recognition of Indian characters is very difficult with compare to English because of its shape variability of the characters.

There are twelve scripts in India and in most of these scripts the number of alphabets (basic and compound characters) is more than 250, which makes keyboard design and subsequent data entry a difficult job. Hence, online recognition of such scripts has a commercial demand. Although a number of studies [16] have been done for offline recognition of a few printed Indian scripts like Devnagari, Bangla, Gurumukhi, Oriya, etc. with commercial level accuracy, but to the best of knowledge no system is commercially available for online recognition of any Indian script.

There is a proliferation of on-line recognizers developed as compared to off-line recognizers. There are two main reasons for this disparity. First, on-line recognizers are easier to build [3], because the order of the pen-strokes is known, as well as timing information and also direction information of writing may be extract. Secondly, handwriting recognition can easily be used for input in handheld or PDA-style computers, where there is no room for a keyboard. Since a recognizer in this use is very visible, this visibility spurs on development.

In this work a new algorithm is proposed for online Bangla handwritten character recognition from stroke. A database of strokes is generated based on the database of characters collected for the experiment. After recognition of the strokes a tree based approach is used for construction of valid character from its constituting strokes. This algorithm is robust against stroke number and order-variations.

The organization of the paper is as follows. In Section 2 the property of Bangla script as well as the pre-processing is described. Section 3 deals with detailed property of stokes in Bangla character and their extraction strategy. The feature extraction technique is described in Section 4 and the classifier is in Section 5. Finally is Section 6 the result and their analysis is given.

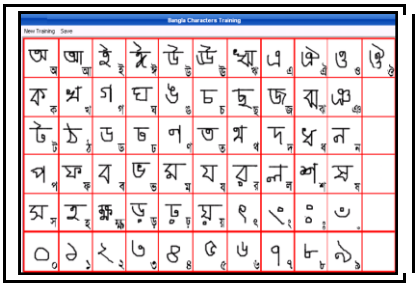
II. BANGLA SCRIPT & PREPROCESSING

Bangla, the second most popular language in India and the fifth most popular language in the world, is an ancient Indo-Aryans language. About 200 million people in the eastern part of Indian subcontinent speak in this language. Bangla script alphabets are used in texts of Bangla, Assamese and Manipuri languages. Also, Bangla is the national language of Bangladesh.

The alphabet of the modern Bangla script consists of 11 vowels and 40 consonants [16]. These characters are called as basic characters. Writing style in Bangla is from left to right and the concept of upper/lower case is absent in this script. It can be seen that most of the characters of Bangla have a horizontal line (Matra) at the upper part. From a statistical analysis on printed document it was noticed that the probability that a Bangla word will have horizontal line is 0.994 [16].

In Bangla script a vowel following a consonant takes a modified shape. Depending on the vowel, its modified shape is placed at the left, right, both left and right, or bottom of the consonant. These modified shapes are called modified characters. A consonant or a vowel following a consonant sometimes takes a compound orthographic shape, which is called as compound character. Compound characters can be combinations of two consonants as well as a consonant and a vowel. Compounding of three or four characters also exists in Bangla. There are about 280 compound characters in Bangla [16]. In this work the recognition of Bangla basic characters are considered.

To get an idea of Bangla basic characters and their variability in handwriting, a set of handwritten Bangla basic characters are shown in Figure 1. This is the form which is also used for collection of isolated characters.



অ	আ	ই	ঈ	উ	ঊ	ঋ	এ	ঐ	ও	ঔ
ক	খ	গ	ঘ	ঙ	চ	ছ	জ	ঝ	ঞ	
ট	ঠ	ড	ঢ	ণ	ত	থ	দ	ধ	ন	
প	ফ	ব	ভ	ম	য	র	ল	শ	ষ	
স	হ	ক্ষ	ড়	ঢ়	য়	ৎ	ং	ঃ	ঁ	
০	১	২	৩	৪	৫	৬	৭	৮	৯	

Fig. 1. Examples of Bangla handwritten characters.

A. Pre-processing

The digitizer output is represented in the format of $P_{i=1}^{M} \in R^{2} x\{0,1\}$, where p_i is the pen position having x-coordinate (x_i) and y-coordinate (y_i) and M is the total number of sample point. For writing Bangla characters, M varies from 14 to 189 for a character. If p_i and p_j are two consecutive pen points, i^{th} point (p_i), is retained with respect to j^{th} point (p_j), if the following condition is satisfied:

$$x^2 + y^2 > m^2 \qquad (1)$$

Where $x = x_i - x_j$ and $y = y_i - y_j$. The parameter m is empirically chosen. M is set to 0; in equation (1) to removes all repeated points.

Analyzing a total of 22,000 Bangla character it was found that, for writing Bangla characters, the number of points varies from 14 (৩) to 189 (ঈ) points. The average number of points in a Bangla character is 72. It was also noted that the character (ঈ) uses the maximum number of points in average and its value is 115. It is closely followed by 'ঊ' (108), 'ক্ষ' (105), & 'আ' (104). The minimum number of points in an average is used by the character 'চ' (47) and is closely followed by 'ব' (49) 'দ' (51).

Smoothing: To remove jitter from the handwritten data, every point (x(t), y(t)) are replaced in the trajectory by the mean value of its neighbors:

$$x'(t) = \frac{x(t-N) + \ldots\ldots + x(t-1) + \alpha x(t) + x(t+1) + \ldots\ldots + x(t+N)}{2N+\alpha}$$

$$y'(t) = \frac{y(t-N) + \ldots\ldots + y(t-1) + \alpha y(t) + y(t+1) + \ldots\ldots + y(t+N)}{2N+\alpha}$$

The parameter '**α**' is based on the angle subtended by the preceding and succeeding curve segment of (x(t), y(t)) and is empirically optimized. This help to avoid smoothing of sharp edges. Here the value for N is taken as 2.

III. STROKE DATA BASE GENERATION

Analyzing of Bangla characters it was found that Bangla characters are formed by combination of one or more basic strokes. The recognition of Bangla script is more difficult compare to roman script due to its large size of character set and compound characters. It gets tougher due to presence of multiple strokes while writing Bangla character. By stroke I mean collection of pen points that are collected between one pen down and pen up (with lifting in between).

The problem of online Bangla handwriting recognition gets more complex due to stroke order variation and variation in number of strokes used to write a character. For example let us consider the character 'আ'. Again it may be seen that 2-6 of the 7 strokes are used for writing the same (as found from statistical analysis of the database). For example see figure 2. So with possible 6 strokes (out of 7) and their order variation makes the recognition process more complicated. Some ways of writing the character 'আ' is shown in figure 3.

The problem is more complex due to stroke order variation. If only 4 strokes are considered to write the character 'আ', they may be again in different order. For example see figure 4.

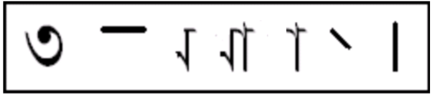


Fig.2. Examples of strokes use to write আ

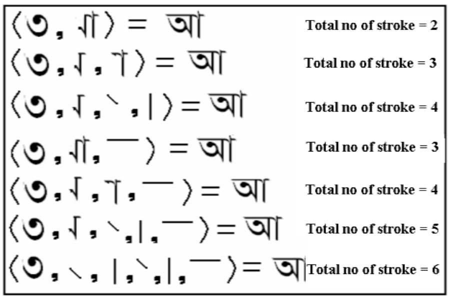


Fig 3: Some of the ways of writing ‘আ’ using some of the possible strokes in some combinations.

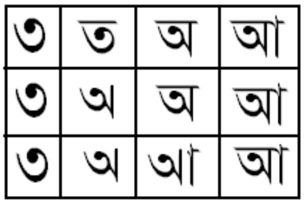


Fig. 4: Example of part of the different stroke-order for a character having four strokes.

A Bangla character may be written with single stroke like ‘ও’ and minimum of two strokes for the character having two disjoint parts like ‘র’. From statistical analysis on the dataset it is found that the minimum number of stroke used to write a Bangla character is 1 and maximum number is 6. It is also seen that ‘এ’, ‘ও’ & ‘९’ have almost written by single stroke where as ঊ has an average of 4 strokes followed by ‘আ’ and rest all characters has average stroke number less then 3. It also found that almost always ‘ঐ’, ‘ঔ’, ‘ত’, ‘ন’, ‘২’, ‘ঃ’ and `ঁ ’ are written by 2 strokes. The average number of stroke per character is 2.2. In the database there are 61 basic characters in Bangla character set, so the total number of stroke in the Bangla Script should have been 135 (app). But from statistical analysis on the present stroke data base it is found that only 59 strokes are enough to represent all basic Bangla Characters. All these basic strokes are shown in figure 5. This is because some of the strokes are common to various characters. For example see table 1. For example the stroke ‘ᐊ is common in the characters ‘ঋ’, ‘ক’, ‘ঝ’, ‘ধ’, ‘ব’, ‘র’. From statistical analysis it is found that 40 out of the 61 characters generally bear Matra (‾). This stroke comprises 24.03% of the stroke database. The second most used stroke is 0 and is used by 8 basic characters and it comprises 8.45% of the stroke database. The top 10 frequent strokes with their percentage of occurrence in the database and the characters in which they occur are detailed in table 1. It is found from the analysis that top 10 frequent strokes comprises 59.76% of the stroke data base and least 20 frequent strokes comprises only 10.24% of the total stroke database. So the character having some of its constituent strokes are common (inter character similarity) are look similar (sometime even the strokes only vary in relative position only) and some time the same character is written with different strokes (intra character variability) makes the recognition process more complicated.

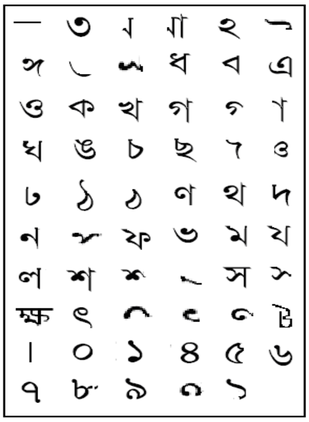


Fig. 5: Basic Stroke of Bangla Script.

Table 1: Most frequent strokes (with percentage) and the characters in which they occur.

Stroke	Occurrence in Database	No of Character having this stroke (Characters)
—	24.01%	40
o	08.45%	8 (‘র’, ‘ড়’, ‘ঢ়’, ‘য়’, ‘o’, ’ ঁ’ং, ‘ঃ)
ᓗ	05.64%	8 (‘ই’, ‘ঈ’, ‘উ’, ‘ঊ’, ‘ঐ’, ‘ঔ’, ‘ট’, ‘ঠ’)
৬	04.99%	6 (‘উ’, ‘ঊ’, ‘জ’, ‘ড’, ‘ড়’, ‘৬’)
ᐊ	03.09%	6 (‘ঋ’, ‘ক’, ‘ঝ’, ‘ধ’, ‘ব’, ‘র’)

৩	02.79%	4 (‘অ’, ‘আ’, ‘ত’, ‘৩’)
এ	02.72%	3 (‘এ’, ‘ঐ’, ‘ঞ’)
া	02.71%	5 (‘অ’, ‘আ’, ‘ঋ’, ‘ঝ’, ‘স’)
য	02.67%	4 (‘য’, ‘য়’, ‘ফ’, ‘ষ’)
২	02.56%	3 (‘ই’, ‘হ’, ‘২’)

A data collection form was prepared for isolated character collection but stroke can’t be collected similarly. This is because normal human are not acquainted with writing strokes individually and secondly if strokes are written individually then the stroke variation (shape, size and number) may not reflect properly. So here the strokes are extracted from their parent characters. To extract the strokes properly from characters for training of the proposed system, the strokes are classified in the following categories:

Major Stroke: A stroke is called Major stroke if it occupies major part of the character. For the character `অ` the major stroke is `৩’.

Minor Stroke: A stroke is called Minor stroke if it occupies minor part of the character. For the character `অ` the minor stroke is `া

Upper Stroke: A stroke is called Upper stroke if it occurs at the upper part of the character. For the character `ই` the upper stroke is ` ’

Lower Stroke: A stroke is called Lower stroke if it occurs at the lower part of the character. For the character `র` the lower stroke is `ο’.

Left Stroke: A stroke is called Left stroke if it occurs at the left part of the character. For the character `ধ` the left stroke is ` ’

Right Stroke: A stroke is called Right stroke if it occurs at the right part of the character. For the character `অ` the right stroke is ` ’.

Matra: A stroke is called Matra if it is a horizontal straight line occurred at the upper part of the character. For the character `অ` the major stroke is `—’.

Depending on the positional information all stroke of a character are classified into above classes. This classification has two important roles in this recognition. As this is a stroke based approach, so training is done on all individual stroke. But this stroke are not collected separately rather they extracted, identified and classified from the character data.

Second important role played by the stroke in time of matching of identified stroke in to characters. This time to overcome the problem of different sequence of stroke order even in case of same character, Major stroke and Matra are find out and they are placed at first and last position respectively in stroke sequence. This minimizes the permutation in tree structure and it gives a greater freedom to consider different confidence value of non major character.

IV. FEATURE EXTRACTION

Any online feature is very much sensitive to writing stroke sequence and size variation. A total of 105 features (90+15) are used for recognition. The features used are (i) Structural features (15) and (ii) Point based feature (90).

4.1 Structural features

Gradient (t_{N+1}):

$$t_{N+1} = n\left(\sum_{i=0}^{N} x_i y_i\right)\left(\sum_{i=0}^{N} x_i\right)\left(\sum_{i=0}^{N} y_i\right)\left(\sum_{i=0}^{N} x^2 - \left(\sum_{i=0}^{N} x\right)^2\right)$$

$$t_{N+2} = \left(\sum_{i=0}^{N} y_i\right)\left(\sum_{i=0}^{N} x^2\right)\left(\sum_{i=0}^{N} x_i\right)\left(\sum_{i=0}^{N} x_i y_i\right)\left(\sum_{i=o}^{N} x_i^2 - \left(\sum_{i=0}^{N} x_i\right)^2\right)$$

Here t_{N+1} and t_{N+2} are the gradient and the intercept in the y-axis of the straight line, constituting by the consecutive 3-points respectively.

Length by Width ratio (t_{N+3}): $t_{N+3} = (\max(x_i) - \min(x_i)) / (\max(y_i) - \min(y_i))$ $\forall$ $i = 0, 1,\ldots,N$ By using this feature the ratio of the length and width of the corresponding stroke is calculated.

Standard Deviation (t_{N+4}): The standard deviation measures the spread of the data about the mean value. It is useful in comparing sets of data which may have the same mean but a different range. Here the deviation of each co-ordinate is calculated with respect to its mean value.

Normalized Start Co-ordinates and End Co-ordinates (t_{N+4}): In this feature only the first and last co-ordinates in the strokes of a character considered. Taking the first and last co-ordinates normalized them and stored them as feature.

Crossing of the lines: Here the co-ordinate position of the crossing of the stroke is stored with itself as shown in figure 6. In this system only first two crossing are considered.

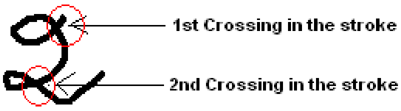
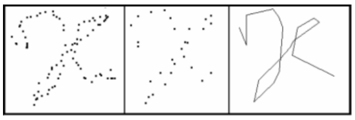


Fig. 6: Crossing points of a stroke.

4.2 Point based feature

The strokes are first normalized in to 30 points. The normalization is done in two stages. First the points are re-sampled to fixed number points and then they are converted from equal time sample to equal distant points. For example see Figure 7. The processed character is transformed into a sequence $t = [t_1,.., t_N]$ of feature vectors $t_i = (t_{i1,}t_{i2,}t_{i3})^T$ [4]. The following features were calculated:



(a) (b) (c)

Fig. 7. Feature extraction from a sample stroke is shown. (a) Original stroke, (b) its normalized 30 points used as feature, (c) the normalized stroke.

Normalized horizontal (t_{i1}) **and vertical** (t_{i2}) **co-ordinates**: $t_{i1}=(x_i-\mu_x)/\sigma_y$ and $t_{i2}=(y_i-\mu_y)/\sigma_y$ are the pen co-ordinates normalized by the sample mean $\mu=\frac{1}{N}\sum_{i=1}^{N}P_i$ and standard deviation, $\sigma=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^2}$ of the character's sample points.

Tangent slope angle (t_{i3})**:** $t_{i3} = \arg((x_{i+1} - x_{i-1}) + j*(y_{i+1} - y_{i-1}))$, with $j^2 = -1$ and "arg " the phase of the complex number above, is an approximation of the tangent slope angle at point *i*.

Thus finally, a feature vector sequence is defined as $t = [t_1\dots\ t_N\ t_{N+1}\ _{\dots.}\ t_{N+15}]$, each vector of it as $t_i = (t_{i1,}t_{i2,}t_{i3})^T$ is obtained. Here the number of points in which the character is normalized is (N) 30. So a total of 105 (30 X 3 [3 for each point] + 15 [15 local features based on Stroke]) features are used.

V. RECOGNITION

In this work, the recognition module has been divided into two parts: (i) Recognition of input strokes and (ii) Construction of valid character from recognized strokes.

5.1 Recognition of input strokes:

Based on the above-normalized features, a Multilayer Perceptron Neural Network based scheme was used for recognition of the strokes [18]. The Multi Layer Perceptron Network (MLP) is, in general, a layered feed-forward network, pictorially represented with a directed acyclic graph. Each node in the graph stands for an artificial neuron of the MLP, and the labels in each directed arc denote the strength of synaptic connection between two neurons and the direction of the signal flow in the MLP. For pattern classification, the number of neurons in the input layer of an MLP is determined by the number of features selected for representing the relevant patterns in the feature space and output layer by the number of classes in which the input data belongs. The neurons in hidden and output layers compute the sigmoidal function on the sum of the products of input values and weight values of the corresponding connections to each neuron.

Training process of an MLP involves tuning the strengths of its synaptic connections so that it can respond appropriately to every input taken from the training set. The number of hidden layers and the number of neurons in a hidden layer required to design an MLP are also determined during its training. Training process incorporates learning ability in an MLP. Generalization ability of an MLP is tested by checking its responses to input patterns which do not belong to the training set.

Back propagation algorithm, which uses patterns of known classes to constitute the training set, represents a supervised learning method. After supplying each training pattern to the MLP, it computes the sum of the squared errors at the output layer and adjusts the weight values of the synaptic connections to minimize the error sum. Weight values are adjusted by propagating the error sum from the output layer to the input layer.

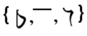
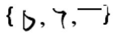
The present work selects a 2-layer perceptron for the handwritten numeral recognition. The number of neurons in input and output layers of the perceptron is set to 105 and 59; respectively since the number features is 105 and the number of possible classes in hand written stroke considered for the present case is 59. The number of hidden units was set to 90, back propagation learning rate and acceleration factor is set to suitable values, based on trial runs. A network of 105-90-59 is thus finally designed.

5.2 Construction of valid character from recognized strokes:

Each character will be constructed with the help of its recognized strokes. To do so, all the probable sequences of strokes are stored in a tree structure that makes a valid character into a database. To build this a database report is generated from the raw data (characters), from which the sequences of strokes of the characters are gotten, that are generally used by people.

5.3 Construction of the Rule base:

The database has been designed using a tree structure to store the possible sequences of strokes of the characters. To store the sequences a stroke is considered as a root.

Figure 8 represent the stroke sequences of ‘ছ’. According to above tree structure, there exist two probable sequences of ‘ছ’.The first sequence is  and second is .

The classifier returns a set of the recognized strokes with their corresponding confidence values. Here it is plain s to consider only the top three choices; number of choices can be extended in future as the system requirement. With these recognized strokes it will be tried to match those sequences with the stored sequence of strokes in the database. When a match will be found then the character recognized as a valid character and all the other combinations will be discarded.

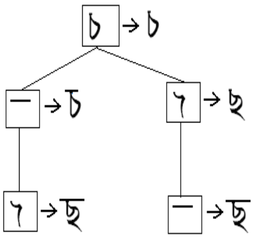


Fig. 8: Sample tree structure.

VI. RESULTS AND DISCUSSION

The experimental evaluation of the above techniques was carried out using isolated Bangla strokes. The data was collected from people of different background. A total of 21372 characters are collected for the experiment. Out of them strokes were extracted and total size of the stroke database used for the training was 30,000. One fourth of the strokes were used for the training of the classifier for the present work and rest will be used for the testing purpose.

6.1 Recognition result on isolated Stroke

The recognition rate of the isolated strokes was found to be 96.85% on the test set.

6.2 Recognition results on Bangla Character.

From the experiment it was found that the overall accuracy of the proposed scheme was 88.23% without rejection. The accuracy improved to 92.9%, if we consider first three top choices of the recognition results. The detail recognition results are given in Table 2.

Maximum error occurred between ‘ম’ and ‘স’ and it is noted that about 1.067% cases they miss-recognized one as the other. Here the difference between the above two characters is that there is a small loop in left bottom side of one of the character. Some times during handwriting people do not give this loop and hence miss-recognition occurs. The next erroneous character is ‘ঘ’. It is misclassified with ‘থ’ and some time with ‘ম’, because difference between ‘ঘ’ with ‘থ’ and ‘ম’ is that ‘থ’ have a small loop at upper left and ম have the loop at lower left other wise they are similar. Sometime writer do not give this small loop and neglect all other very small difference. ‘ঘ’ is miss-recognized as ‘থ’ in 5 cases and miss-recognized as ‘ম’ in 8 cases. For details see Table 3.

Table 2: Top three recognition accuracy are shown.

Choices from top	Recognition rate
1	88.23%
2	91.57%
3	92.93%

Table 3: Maximum confusion pair arte shown.

Character	Miss-Recognized	
ঘ	থ(8%)	ম(13%)
ম	ঘ(1%)	স(1%)
ঢ	চ(1%)	
ন	ণ(11%)	

This work describes a novel system of 0nline Bangla Handwriting Recognition. This work is on a preliminary stage and it is hoped that the result will improve when more and combinations of stroke will be added to the rule base. After completion of the first stage of isolated character recognition the system will be modified for unconstrained Bangla online text recognition. I think this work will be helpful for research in Bangla online handwriting recognition. However it not only help the development of Bangla script but also other Indic script which have a lot of similarity with Bangla script with adequate modification.

REFERENCES

[1] I. Guyon, M. Schenkel, and J. Denker, "Overview and Synthesis of On-Line Cursive Handwriting Recognition Techniques", Handbook of Character Recognition and Document Image Analysis, 183 (1997).

[2] E. J. Bellagarda, J. R. Bellagarda, D. Nahamoo and N. S. Nathan, "A probabilistic Framework for Online Handwriting Recognition," 3rd IWFHR, 234 (1993).

[3] R. Plamondon and S. N. Srihari, "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey". IEEE PAMI, **22**, 84 (2000).

[4] C. Bahlmann and H. Burkhardt, "The Writer Independent Online Handwriting Recognition System frog on hand and Cluster Generative Statistical Dynamic Time Warping", IEEE PAMI, **26**, 1 (2004).

[5] C. C. Tappert. C. Y. Suen, and T. Wakahara. "The State of the Art in On-Line Handwriting Recognition," IEEE PAMI, 12, 179 (1990).

[6] S. Jaeger, C. L. Liu and M. Nakagawa: "The state of the art in Japanese online handwriting recognition compared to techniques in western handwriting recognition", IJDAR, **6**, 75 (2003).

[7] C. L. Liu, S. Jaeger and M. Nakagawa: "On-Line Recognition of Chinese Characters: the State of the Art", IEEE PAMI, **26**, 198 (2004).

[8] S. Jaeger, S. Manke, J. Reichert, A. Waibel, "Online handwriting recognition the Npen++ recognizer", IJDAR, **8**, 169 (2001).

[9] Zhen-Long BAI and Qiang HUO, "A Study On the Use of 8-Directional Features For Online Handwritten Chinese Character Recognition", ICDAR, (2005).

[10] V. Bansal and R. M. K. Sinha, "On how to describe Shapes of Devanagari Characters and Use them for Recognition." ICDAR (1999).

[11] S.D. Connell and A. K. Jain, "Template-based Online Character Recognition", PR, **34**, 1 (2001).

[12] U. Garain, B. B. Chaudhuri, and T. Pal, "Online Handwritten Indian Script Recognition: A Human Motor Function based Framework", ICPR (2002).

[13] S. D. Connell, R. M. K. Sinha, and A. K. Jain, "Recognition of Unconstrained Online Devnagari Characters", ICPR (2000).

[14] N. Joshi, G. Sita, A. G. Ramakrishnan, V. Deepu, Sriganesh Madhvanath: "Machine Recognition of Online Handwritten Devanagari Characters", ICDAR (2005).

[15] K. Roy, N. Sharma, T. Pal and U. Pal, "Online Bangla Handwriting Recognition System", ICAPR (2007).

[16] B.B. Chaudhuri and U. Pal, "A Complete Printed Bangla System," PR, **31**, 531 (1998).

[17] U. Pal and B. B. Chaudhuri, "Indian script character recognition: a survey", PR, **9**, 1887 (2004).

[18] K. Roy, C. Chaudhuri, M. Kundu, M. Nasipuri and D. K. Basu, "Comparison of the Multilayer perceptron and the nearest neighbor classifier for handwritten digit recognition", JISE, **21**, 1245 (2005).

[19] U. Bhattacharya, B. K. Gupta and S. K. Parui, "Direction Code Based Features for Recognition of Online Handwritten Characters of Bangla", ICDAR (2007).