Handwritten Devnagari Digit Recognition using Fusion of Global and Local Features

Pratibha Singh

Institute of Engineering and Technology, DAVV Khandwa Road Indore(M.P.), India Ajay Verma Institute of Engineering and Technology, DAVV Khandwa Road Indore(M.P.), India Narendra S. Chaudhari Indian Institute of Technology Khandwa Road Indore(M.P.), India

ABSTRACT

We give our formulation for a ten class classification of handwritten Hindi digit recognition. Automatic Recognition of Handwritten Devnagri Numerals is a difficult task, because of the variability in writing style; pen used for writing and the color of handwriting, unlikely the printed character. Furthermore, Hindi Digit can be drawn in different sizes. Therefore, a robust offline Hindi handwritten recognition system has to account for all of these factors. Hence we have chosen a combination of global and local features. The global features are the structural features like endpoint, crosspoint, centroid of the loop, u shaped structure, C shaped structure and inverted C shaped structure. The local set of features combine the distance of thinned image from geometric centroid calculated zone-wise and histogram based features calculated zone-wise. Variability in writing style is taken care by size normalization and normalization to constant thickness as preprocessing a step before feature extraction. We used an Artificial Neural Network as classifier for recognition. Our method results in average correct rate of 95% or better. The combination of local and global features results in reduced confusion value..

General Terms

Pattern Recognition.

Keywords

Features, ANN, Structural feature, neuron, Histogram

1. INTRODUCTION

The need to recognize the handwritten text is challenging problem not only from the perspective of behavioral biometrics but also in the context of pattern recognition. Although many pieces of work had been done for recognition of roman script, only few attempts have been tried for Indian scripts. Hindi is the official and widely spoken language of India which is written and encoded using Devnagri script. There are two fundamental approaches to character recognition: Feature based classification and Template matching. Template matching based approach is sensitive to size and style variation; therefore we go for the first one, *i.e.* feature based classification.

Optical Character Recognition is a process of automatic recognition of different characters from a document image. The task can broadly be separated into two categories: the recognition of machine printed data and the recognition of handwritten data. Machine printed characters are uniform in size, position and pitch for any given font. In contrast, handwritten characters are non-uniform; they can be written in many different styles and sizes by different writers and at different times even by the same writer. Recognizing handwritten numerals is an important area of research because of its various application potentials. Automating bank cheque processing, postal mail sorting, job application form sorting, automatic scoring of tests containing multiple choice questions, are amongst a few applications where numeral recognition is necessary.

The problem of numeral recognition has been studied for decades and many methods have been proposed such as template matching, dynamic programming, hidden Markov modeling, neural network, expert system and combinations of all these techniques. Feature extraction [1] plays a vital role in image processing system in general and character recognition system in particular. First research report on Devnagri script was published in 1977[2] but not much work has been reported for next two decades. Some research has been reported recently on handwritten Devnagri characters. Hanmndlu and Murthy [3] used fuzzy model based recognition for handwritten Hindi numerals, where they used Normalized distance as feature for individual boxes and obtained 92.67% accuracy. Ramteke et al [4] proposed handwritten isolated Marathi numeral recognition scheme based on invariant moments. They used Gaussian distribution function for classification and obtained 87 % accuracy. Bhattacharaya et al [5] proposed a multilayer perceptron (MLP) neural network based classification approach for Devnagri numerals and obtained 91.28 % results. They used multiresolution features based on wavelet transforms. Sharma et al [6] proposed quadratic classifier based approach for Devnagri character recognition and obtained 98.86 % accuracy for numerals and 80.36 % for characters. S. Arora [10,11] presented the recognition of Devnagri recognition method using shadow features, view based features, longest run feature, and chain code based features. The decision of classifiers is combined using majority voting and weighted majority weighting. In the experiments produced by Impedovo [13] the most valuable spread of features in diverse zones is presented in the form of membership functions for the respective zones for handwritten digit recognition. Dimauro [14] presented an approach for numeral recognition by weighting the local decisions of respective zone. The

optimal weight vector for the combination of respective zones is obtained using genetic algorithm.

This has motivated us to design a simple and robust algorithm for handwritten numerals recognition, which is independent of size, slant, ink, and writing style. In this paper, seven different categories of structural features are combined to obtain high degree of accuracy for handwritten numeral recognition. We have identified potential feature sets including end point, cross point, loop, four Maximum profile features and three Water Reservoir based features.

There are 10 numerals. The sample set of these numerals is as shown below:-



Figure 1 Data sample of collected Hindi Numerals

The recognition of the above handwritten Devnagri numerals is the main objective of the research

2. PROPOSED APPROACH

The figure 2 gives the design cycle for the proposed approach



Figure 2 A general diagram for digit recognition

2.1 Input data acquisition

The generalizability of the results from any OCR endeavor depends a lot on the sample database used. The objective in data collection is to obtain a set of handwritten samples of Devnagari numerals that capture effectively the huge variations in handwriting between and within writers so that the database is truly representative of the variety encountered in the field. This is necessary for getting a realistic assessment of the power of the proposed methodology. For creation of our datasets of numeral samples, the following criteria were specified:

a) The persons writing the numbers were free to use different quality pens, different ink color etc.

- b) They were required to write the numerals in the specified grids while ensuring that the numerals did not touch the grid lines.
- c) Each person was asked to write 0-9 (in Devnagari Script). 3,000 samples were collected using the above approach.

We have scanned the sample pages through a HP 3050 series scanner of resolution 300 dpi. The images stored in grey scale .png format, which is converted to two tone binary image using OSTU [8] binarization algorithm. For each digit, 300 samples were being collected, out of the 300 samples, first 220 samples were selected whose features were computed and stored in the database and the remaining 80 samples were used as test images. Care was taken to ensure that test cases were from persons different from those who provided the database samples. This is necessary for a more realistic assessment of the recognition capability of the proposed approach.

2.2 Preprocessing

In the scanning process, some distortion in images may be introduced due to writing of different peoples. Some preprocessing steps are performed for rectification of distorted images, improving the quality of images. The outputs of the following preprocessing steps are shown in figure 3. The main objectives of pre-processing are:

- Noise removal: Noise removal refers to removal of any unwanted or insignificant bit pattern in the image which is simply acting like a noise.
- b) Normalization: Handwriting produces variability in size of written digits. This leads to the need of scaling the digits size within the image to a standard size, as this may lead to better recognition accuracy. We normalized the size of digit within the image and also translated it to a specific position by the centralization on 250*250.
 - c) Skeletonization or thinning: In more practical terms, thinline representations of patterns would be more amenable to extraction of features such as end points, junction points and connections among the components. For thinning algorithm to be really effective, it should ideally compress data, retain significant features of the pattern and eliminate local noise without introducing distortions of its own. Thinning is normally applied to binary images and produces another binary image as output. As the skeletonization is always performed on white portion of the image, the image first is inverted. Later skeleton function is performed on image.
 - d) Spur removal: After performing skeletonization, the result shows that there is still some noise present in the image in the form of spur. To remove this spur, spur removal function is preformed.
 - e) Normalization to constant Line Width: The data collected by various people consist of varying widths; this is because some may choose to write with pointed tip pen and others may use a broader tip. For our histogram based feature set extraction, the stroke width normalization is performed which accepts the skeleton of the image and produces the uniformly thick output of the character image. This step is shown in figure 4.



Figure 3 Outputs of preprocessing steps of character, (a)Gray Image (b) Black & white image (c) Skeletonized Image (d) Image after spur Removal



Figure 4 Constant width representation of the character

2.3 Feature extraction

In feature extraction stage each character is represented as a feature vector, which becomes its identity. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements. Feature extraction methods are based on three types of features:

- a) Statistical features: Statistical features are derived from the statistical distribution of pixels and describe the characteristic measurements of the pattern
- b) Structural features: It describes geometrical and topological characteristics of a pattern by representing its global and local properties
- c) Global transformation and moments: Global transformation technique transforms the pixel representation to a more compact form. This reduces the dimensionality of the feature vector and provides feature invariants to global deformation like translation, and rotation.

In this paper, features used are same as that used in [15] and [16]. However we have combined the global (structural) features used in [16] with the local (histogram) features used in [15]. In this paper we report better recognition in terms of "correct rate" by combining global features and local features. The feature sets are:

Set 1: These are considered as the potential features; they are based on the structural pattern. Structural analysis includes features such as junction points, end points, loops etc. These features are unique to each numeral. The junction point and end point form the characteristic points. For example, numeral 6 has one junction point and one end point and numeral 8 may have single junction point or two junction points. It depends on how the numeral is written. The maximum profile distance feature and structures like U, C and inverted C (for example the numeral 3) are the additional structural features used for the numeral recognition. The characteristic points are extracted from the character image, which are described as:

• *End point*: Each pixel of a skeleton, resulting from the thinning process, represents the connectivity between pixels and its eight neighbors. End point has the property of one point connectivity. If any pixel has one point connectivity that means it is end point.



Figure 5 Extraction of End point

The bounding box of character is divided into 9 zones and the zone containing the end point is represented as 1 and the rest as 0; so the database contains 9 values corresponding to one end point.

• Junction point: It is also defined as Cross point. It has the property of three point connectivity. A junction point is pixel which is surrounded by two or more pixels in its neighbor.



The bounding box is divided into 9 zones and the zone containing the cross point is represented as 1 and the rest as 0. The database contains 9 values corresponding to cross point.

• *Loop*: The loop is identified by their property of closed bound. After filling holes in the binary image, loop can be identified.



Figure 7 Extraction of loop

After the detection of loop, centroid is calculated for the loop and stored in the database of feature set 1.

• *Maximum profile distance*: After fitting the bounding box of each numeral on 250*250 image size, their profiles are computed in four directions (top, bottom, right and left). Thus the maximum profile distance from the bounding box edge is obtained in four directions, the profile feature computations are illustrated in Figure 8.



Figure 8 Method of obtaining maximum profile distance

- U' shape structure: This structure can be considered as water reservoir based structure. The water reservoir based principle is that, if water is poured from one side of a component, the cavity regions of the component where water will be stored are considered as reservoirs. The width and height of 'U' structure is calculated and is included in database.
- 'C' shape structure: This structure is also based on water reservoir but here the principle is that, Right Reservoir can be obtained when water is poured from right of the component then the structure we get is like C. The width and height of 'C' structure is calculated and is included in database.

These above six values are included in database as set 1.

Set 2: These are the local features which are obtained by dividing the image into zones. We extracted three different types of features: the one is centroid distance feature and the other two are extracted from constant width image of the character.

- *Centroid Distance feature*: In this feature centroid of the image is calculated and the average distance from centroid to each and every point is calculated. Further the image is segmented into four zones and then individually the centroid of all four zones is calculated and their average distances in the respective zones forms the feature set.
- *The Number of peaks in Histogram*: A Histogram is a Graphical representation, showing the visual impression of distribution of data. The figure 9 shows the histogram of Devnagri numeral three..



Figure 9 Horizontal and vertical histogram of numeral 'teen'

The figure 10 shows that number of peaks in Horizontal Histogram is two and that in Vertical Histogram is three.



Figure 10 The number of peaks and the length of peak of histogram

 Sorted Density in Histogram: Histogram so extracted in the previous feature is divided into several zones and density per zone is calculated. Further the density calculated per zone is sorted to get top 9 densities so that a fixed feature length can be obtained. Vertical Histogram divided into vertical zones and horizontal histogram is divided into horizontal zones.



Figure 11 Horizontal and vertical zoning of histogram

3. CLASSIFICATION

Classification is the important stage for numeral character recognition. The extracted features such as end points, junction points and loop and pattern are used to identify a numeral. Pattern is used for matching with reference database of numeral 0 to 9 digits. A neural network-based classification scheme is designed for this task.

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. For each neuron in this network, it outputs an activation value as a function of its net input through an activation function. Here, a unipolar sigmoid function is selected as the activation function.

The network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input/target pairs are used, in this supervised learning, to train a network.



Figure 12 A Neural Network for Classifying Character Categories

The Network designed for our each set feature is 3 layer feedforward neural network having input nodes equal to the number of features. For feature set 1 there are 30 neurons in the input layer and 18 neurons in the hidden layer. For feature set 2 we have 25 neurons in the input layer and 18 neurons in hidden layer. For combined set, we have 55 neurons in the International Journal of Computer Applications (0975 – 8887) Volume 89 – No.1, March 2014

input layer and 30 neuron in the hidden layer. For every the network, 10 neurons in output layer (equal to number of classes) are used. Back propagation Levenberg-Marquardt algorithm with momentum is used for training. The transfer function for input and hidden layer is tansigmoid. The transfer function of output layer was logsigmoid. The learning rate of 0.1 is used for training.

4. RESULTS

The handwritten numeral character test data consists of 80 samples for each class, while 220 samples per class are used for training. The various types of large variability of handwritten numeral characters test samples are taken.

The correct rate for each digit is shown in given in Table I and confusion values for set 1, set 2 and their combination is shown in table II, III and IV. The correct rate is defined as

Correct Rate = $\frac{correctly \ classified \ samples}{classified \ samples}$

Confusion value= fraction of misclassified samples

Feature	0	1	2	3	4	5	6	7	8	9
set										
Set1	1	0.9975	0.9624	0.9549	0.9925	0.9875	0.9925	0.9887	0.9912	0.9962
Set 2	0.9013	0.9725	0.93	0.9325	0.9175	0.9225	0.9313	0.9363	0.9875	0.9375
Set1+Set2	0.9987	0.9987	0.9724	0.9724	0.9925	0.9875	0.9912	0.9987	0.9987	0.9975

Table 2 Confusion matrix for set1											
Digit /	0	1	2	3	4	5	6	7	8	9	
Classified as											
0	80	0	0	0	0	0	0	0	0	0	
1	0	79	0	0	0	0	0	1	0	0	
2	0	2	60	10	1	0	3	0	4	0	
3	0	0	10	62	1	4	1	1	1	0	
4	0	0	0	0	78	0	1	1	0	0	
5	0	0	0	4	1	73	2	0	0	0	
6	1	2	1	2	0	0	73	0	1	0	
7	0	0	0	1	1	1	6	68	1	2	
8	0	0	0	0	0	0	0	9	71	0	
9	1	1	2	1	1	1	1	1	9	62	
Confusion value: 0.0950											

Table 3 Confusion matrix for set2

Digit /	0	1	2	3	4	5	6	7	8	9
Classified as										
0	68	0	1	0	6	0	1	0	2	2
1	13	58	0	0	2	5	0	0	0	2
2	16	0	41	8	0	9	4	0	0	2
3	12	2	10	43	0	10	1	1	1	0
4	19	1	1	2	39	0	6	9	0	3
5	11	3	0	4	0	52	2	1	2	5
6	13	0	2	2	7	1	46	6	0	3
7	15	0	1	1	3	1	3	52	0	4
8	1	0	0	0	2	0	1	0	76	0
9	8	1	0	0	6	0	2	1	0	62
Confusion value: 0.3287										

Digit /	0	1	2	3	4	5	6	7	8	9
Classified as										
0		-			-	~				
0	80	0	0	0	0	0	0	0	0	0
1	0	80	0	0	0	0	0	0	0	0
2	5	0	65	6	1	1	1	0	1	0
3	6	1	7	64	0	2	0	0	0	0
4	4	0	0	0	75	1	0	0	0	0
5	2	0	0	2	0	76	0	0	0	0
6	4	1	1	1	0	1	72	0	0	0
7	0	0	0	0	1	1	6	70	1	1
8	1	0	0	0	0	0	1	7	71	0
9	1	1	1	1	1	1	1	1	9	63
Confusion value: 0.0725										

Table 4 Confusion matrix for set1+set2

Comparison of correct rates



Figure 13 Comparison of correct rate

5. CONCLUSIONS

In this paper, we use a combination of global and local features for recognition of handwritten numeral characters. Global features are extracted from special points like end point, junction point, loop, special structure 'U' and 'C'. Further profile is calculated for all the four directions and the maximum distance is obtained from the boundary. In local features the image portioning is done and the features obtained were based on distance from the centroid in each of the zones and horizontal and vertical histogram. The variation of handwriting due to various pen thicknesses is counteracted by obtaining skeleton and then converting into constant thickness image. These features were analyzed using neural network classifier. This approach is a very powerful technique in the field of pattern recognition, when noisy and distorted patterns arise. For numeral characters 0, our proposed system results in 100% correct rate. The confusion value for the combination of set 1 and set 2 is reduced to 7.25 %. Further, as demonstrated in section 4, our combination of global and local features results in better "correct rate".

6. REFERENCES

- Q. Due Trier, A.K. Jain, T. Taxt, "Feature Extraction Methods for Character Recognition: A Survey", *Pattern Recognition*, 1996, 29(4), pp. 641-662
- [2] I.K. Sethi and B. Chatterjee, "Machine Recognition of constrained Hand printed Devnagari", *Pattern Recognition*, Vol. 9, pp. 69-75, 1977
- [3] M. Hanmandlu and O.V. Ramana Murthy, "Fuzzy Model Based Recognition of Handwritten Hindi Numerals", *Pattern Recognition*, Volume 40 Issue 6, June, 2007 Elsevier Science Inc pp 1840-1854.
- [4] R. J. Ramteke, P. D. Borkar, S. C. Mehrotra, "Recognition of Isolated Marathi Handwritten Numerals: An Invariant Moment Approach", at Proc. of International Conference on Cognition and Recognition (ICCR 2005), Mysore, (Karnataka), India, pp. 482 - 489 on 22-23 Dec. 2005.

- [5] U. Bhattacharya, B. B. Chaudhuri, R. Ghosh and M. Ghosh, "On Recognition of Handwritten Devnagari Numerals", In Proc. of the Workshop on Learning Algorithms for Pattern Recognition (in conjunction with the 18th Australian Joint Conference on Artificial Intelligence), Sydney, pp.1-7, 2005.
- [6] N. Sharma, U. Pal, F. Kimura and S. Pal "Recognition of Off-Line Handwritten Devnagari Characters Using Quadratic Classifier", Indian confrence on computer vision, graphics and image processing 2006, LNCS 4338, pp. 805 – 816.
- [7] S.V. Rajashekararadhya and P.V. Ranjan, "Efficient Zone based feature extraction method for handwritten numeral recognition of four popular south Indian scripts" *Journal of Theoretical and Applied Information Technology*, 2005-2008, vol4 no12.
- [8] N.Otsu, A threshold selection method from gray level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1), pp. 62-66,1979
- [9] C.-L. Liu, K. Nakashima, H. Sako, H. Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, *Pattern Recognition*, 37(2): pp 265-279, 2004.
- [10] S. Arora, D. Bhattacharjee, M. Nasipuri, M.Kundu, D.K. Basu, "Application of Statistical features in Handwritten Devnagari Character Recognition", *International Journal* of Recent Trends in Engineering, IJRTE Nov 2009 pp 40-42.
- [11] Sandhya Arora, Debotosh Bhattacharjee, Mita Nasipuri, L. Malik , M. Kundu and D. K. Basu, "Performance

Comparison of SVM and ANN for Handwritten Devnagari Character Recognition", *IJCSI International Journal of Computer Science* Issues, Vol. 7, Issue 3, No 6, May 2010 pp 18-25.

- [12] Chun Lei He, Louisa Lam, Ching Y. Suen "Automatic Discrimination between Confusing Classes with Writing Styles Verification in Arabic Handwritten Numeral Recognition", proceedings of IEEE International Conference on Pattern Recognition 23-26 Aug. 2010 pp 2045 – 2048, Istanbul.
- [13] S. Impedovo, R. Modugno, G. Pirlo "Membership Functions for Zoning-based Recognition of Handwritten Digits", IEEE International Conference on Pattern Recognition, Istanbul, Turkey August 23-August 26 2010 pp 1876 - 1879.
- [14] G. Dimauro, S. Impedovo, R. Modugno, G. Pirlo, "Numeral Recognition by Weighting Local Decisions", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), vol. 2, pp.1070 Edinburgh, Scotland.
- [15] P. Singh, J Sabharwal, A. Verma, N. S. Chaudhari, "An Efficient method for the Devnagri handwritten vowel recognition", Accepted for Indian International conference on Artificial Intelligence, IICAI 2011 to be held in Dec.14-16, at Bangalore, India.
- [16] P. Singh, S. Gulani, A. Verma, N. S. Chaudhari, "An Intelligent Network for handwritten Devnagri Digit recognition using Structural features", Accepted for Indian International conference on Artificial Intelligence, IICAI 2011 to be held in Dec.14-16, at Bangalore, India.