Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ZONE BASED RELATIVE DENSITY FEATURE EXTRACTION ALGORITHM FOR UNCONSTRAINED HANDWRITTEN NUMERAL RECOGNITION

K.N. SARAVANAN[1], [2]Dr. R. ANITHA

[1]Assistant Professor, Christ University, Bangalore, Karnataka, India

[2]Prof. & Director, Muthayammal Engg. College, Rasipuram,Tamilnadu, India

E-mail:[1]saravanan.kn@christuniversity.in,[2]aniraniraj@rediffmail.com

ABSTRACT

The recognition of handwritten digit recognition has been a challenging problem among the researchers for few decades. This paper proposes a relative density feature extraction algorithm for recognizing unconstrained single connected handwritten numerals independent of the languages. The proposed method consists of four phases, namely, image enhancement (dilation), representation (zone based), feature extraction (relative density) and recognition (minimum distance classifier). The handwritten numerals must be enhanced with dilation, in order to connect the broken digits. After enhancement, the dilated binary images can be represented as a mid-point aspect ratio class interval values. The minimum distance classifier technique has been used to recognize the given numerals. The method yielded a satisfactory recognition rate of 92.85%, 99.28%, 98.95%, 98.72%, 99.29%, and 99.48% for Latin, Assamese, Devanagari, Manipuri, Malayalam and Oriya Handwritten Numerals respectively.

Keywords: *Handwritten digit recognition, Feature extraction, Zoning, Minimum distance classifier, K-NN classifier.*

1. INTRODUCTION

A numeral recognition system has a variety of commercial and practical applications in postal automation, bank cheque processing, automatic data entry, reading aid for the blinds, vehicle number plate recognition etc. The challenge of building a numeral recognition system which can match the human competence provides a strong motivation for research in this field. The recognition of handwritten numerals, characters, symbols and multifont numerals by computer has been a topic of intense research for almost five decades. A variety of algorithms, combination of features and feature sets have been adopted to try to capture essential information from handwritten numerals. Using different feature sets, and recognition algorithms researchers have obtained accuracy rate up to 99.10% ±0.40 for numerals of their respective languages, but the feature sets and recognition algorithms may not work for other languages because of variant in shapes, orientation etc.[1] The aim of the proposed research work is to recognize the numerals for language independence. The organization of the paper is as follows. In section 2 a brief review of earlier work and in section 3 database collections and pre-processing of handwritten digits. In section 4 representations of the handwritten digits and in section 5 and 6 zone based feature identification followed by feature selection for zone based representation has been described respectively. The proposed methodology has been described in section 7. The experimental results are reported in section 8 and the conclusion of the work is given in section 9.

2. A BRIEF REVIEW OF EARLIER WORK

The strategy used for pattern recognition, especially for handwritten digit recognition can be broadly classified into three categories, namely, Statistical approach , Syntactical approach, Hybrid approach. In statistical approach, a pattern is represented as a vector of an ordered, fixed length list of numeric features. For example, the outer densities of pixels for each of the direction are computed in four directions, namely, top, bottom, left and right. In structural or syntactical approach, a pattern is represented as a set of shapes of an unordered, variable length list of geometric features of mixed type. For example, shapes include end point, fork point and cross point. In hybrid approach, these two approaches are combined for representation of numerals and utilizing them for classification of unknown digits. For example, the Euler number was used as a pre-classification step and statistical features were used for classification. In

Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the following paragraphs, the major works reported in the literature about the three approaches, their recognition rate or error rate and classification algorithms are briefly described. K KAMATA et.al, [2] used structural approach and proposed fourteen pattern primitives as a feature set and primitive tree for classification and has achieved 96% recognition accuracy using 300 handwritten numerals. Keiji YAMADA, Hiroyuki KAMI, Jun TSUKUMO [3], used decision theoretical approach, they used three neural networks, first, globally connected neural network into which a grey level character image is input, second, locally connected neural network into which a grey level character image is input and third, network into which contour features for a character are input and obtained 99.12%, 99.14% for 40 and 100 hidden units respectively. Christine Nadal, Raymond Legault and Ching Y. Suel [4] presented two methods, one classifies samples based on structural features extracted from their skeletons and the other makes use of their contours, and has achieved recognition rate 84.85%. TUAN A. MAI and CHING Y. SUEN, 1990 [5], the computer generated feature is the primary feature used for the coarse classification of numerals, and the secondary features originates from human perception. This study has 16985 samples, 8500 were used in training, the remaining 8485 used in testing, from a training set 407 features were established, recognition result based on inference method and structural method 96.4%, 79.9% respectively. This method works for numerals of English language and thinning algorithms is a mandatory one. The database they have collected from U.S. Postal Services at different locations in the United States. J. T Lin and R.M. Inigo [6] used Back Propagation Neural Network with different structures and images are normalized to size 19 * 19. A lucid scope has been given in the paper as, increasing training numerals the recognition rate would proportionally increase. CHING Y. SUEN, CHRISTINE NADAL et.[7], four algorithms developed independently, three standard sets of data were prepared, each containing 2000 digits (200 of each class). Two of these sets, labelled A and B, are used for training and the third labelled T is used for testing. In the first method, the skeleton of a character pattern is decomposed into parts called branches. The pattern is then classified according to the features extracted from these branches along with the relations which exist among the parts. The recognized rate of 86.05% was achieved. In the second method, the primitives selected to represent the handwritten numerals are line segments, convex polygons and loops. The recognized rate 93.10% was achieved with this method. It has been mentioned that the Expert2 method is a time consuming process, it was considered inadvisable and unnecessary to use the method as the sole classifier. In the third method, the primary features are endpoint, fork point and cross point by the computer and the secondary features are obtained based on the human perception, the recognition rate of 92.95% was achieved. In the fourth method, features are extracted from the contours of the digits, a tree classifier is used and the recognition rate of 93.90% was achieved. Jun Cao, M. Shridhar, F. Kimura, M. Ahmadi [8], each image is divided into rows and columns, called zones. In each zone, a local histogram of the chain code is calculated. The feature vector is composed of these local histograms. They concluded that the neural net classifier outperforms the statistical classifier when the feature vector size is small. Both classifiers are effective in recognizing handwritten numerals with very low error rates and low rejection rates. Heutte L, Moreau J.V., Plessis B, Plagnaud J.L., Lecourtier Y [9], used four feature extraction families, first, concavity measurements, second, horizontal and vertical projections, third, polygonization of both the internal and the external numeral boundaries and fourth feature includes top, bottom, left and right extrema of the numeral. The system was tested on 100,000 digits extracted from the NIST database. 50,000 digits were used for the learning stage and 50,000 digits for the recognition rate. The recognition rate has been 98.05%. Yi Lu, Steven Schlosser, Michael Janeczko [10], used Fourier descriptor to represent the handwritten digits and proposed five Fourier descriptor representations namely, SB, GCT, ERIM1, ERIM2 and DFT and have achieved 99.8%, 98.6%, 99.2%, 98.8% and 99.5% recognition rate respectively. They have also mentioned, a Fourier Descriptor based digit classifier can serve as a pre-classifier in a digit recognition system. Takahiko KAWATANI [11], used four feature extraction families, first, concavity measurements, second, horizontal and vertical projections, third, polygonization of both the internal and the external numeral boundaries and fourth feature includes top, bottom, left and right extrema of the numeral. The recognition rate has been 98.05%. Feng Pan, Mike Keane [12], proposed a new set of aspect invariant moments, which are suitable for neural networks. Their experimental results have proved high recognition rate of 98.73% and low substitution rate of 1.06%. M.H. Shirali-Shahreza e.t.al [13], designed 32 segment bar mask for shadow coding Arabic Numerals irrespective of size and translation and obtained recognition accuracy of 97.8%. Jianming Hu and Hong Yan [14] proposed structural method for describing both printed and handwritten characters. The additional features considered are direction points (D) to characterize the curve changes in horizontal and vertical directions and the bend points (B) are used to detect the curvature changes of a curve in one direction. They have used

Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

102 prototypes for characters and recognition rate of 97.08% was achieved, using thinning algorithms as the pre-processing method. Thien M. Ha and Horst Bunke [15] proposed perturbation approach, reversing an input image to one of its standard forms, CEDAR and NIST database was used which gave a result of 99.09%, 99.54% recognition rate respectively. Xuefang Zhu [16] grouped three different categories of primary features namely, boundary distances in a segment, pixel densities in a segment and line distances from centroid in a segment. To reduce the percentage of the error rate, voting system was used to obtain the final result. Alceu de S. Britto Jr et. al[17] proposed 10 column and 10 row based zoning scheme without making the features size invariant. Their experiments have shown that HMMs can provide high recognition performance 98% close to those provided by the use of Neural Networks (99%). B.V. Dhandra et. al [18] proposed multifont numeral recognition without thinning, using four directional density of pixels. 99.78% of accuracy was achieved and average time required for execution per numeral was found to be 0.0160 seconds using Minimum Distance Classifier. Binu P. Chacko, Babu Anto P [19] used both structural and statistical features and the recognition rate obtained was 93.3% and 95.7% respectively. They have applied thinning algorithm to extract features and represented the image in a 4 * 4 grid. Ying Wen, Pengfei Shi [20] proposed improved LDA and Bhattacharyya distance based classifier for numeral recognition. S.V. Rajashekararadhya, P. Vanaja Ranjan [21] proposed zone and distance metric based feature extraction techniques and computed the distance between the image centroid to each pixel present in the zone. They have obtained 50 features and used support vector machine for classification. They used normalization and thinning during pre-processing stage, the recognition rate 97.25% for Kannada numerals was obtained. S. Impedovo, et.al [22] addresses the problem of selecting the feature membership function for zoning based classification. Several membership functions have been considered, based on abstract level, ranked level and measurement level weighting models and they have found, on average, the exponential membership functions seems to provide the best results, the recognition rate being 85.26%. Mahesh Jangid et.al [23] used two types of features, namely, zoning density and background directional distribution with various zones 4 *4, 5 * 5 and 6 * 6 and obtained 98.76%, 98.91% and 98.54% recognition rate respectively. D. Impedovo, G. Pirlo [24] proposed a multi-objective genetic algorithm with optimal number of zones along with the optimal zones, defined through Voronoi diagrams and obtained a minimum error rate of 6%. Md. Musfiqur Rahman Sazal [25] et al, investigated a feature learning based approach by the deep belief network (DBN) for handwritten Bangle numerals using ISI database and obtained the average accuracy of 91.30%. Ravindra S. Hegadi, Parshuram M. Kamble[26] used Multilayer feed forward neural network for classification and the image was enhanced using dilation and overall recognition rate is 97%. From the literature it reveals that there are methods which are efficient in the recognition of numerals but they requires thinning operation and pruning operation (removal of parasitic components) and also the feature sets extracted from the image are applicable only for the particular language. The challenge is to develop a method that removes size restrictions and language independent. The present study aims at producing a numeral recognition system, which could have numerals of any size, shape and language with reasonable amount of space and time.

3. DATABASE COLLECTION AND PRE-PROCESSING OF HANDWRITTEN DIGITS

In this work, the MNIST standard database comprising of English numerals has been used for both training and testing. It has separate training set and testing set data. Total digits of 58631 used for training purpose and 9621 handwritten samples were used for testing purpose. To validate the numerals available in the database, enhancement and size limits in terms of tolerance limits have been applied. In the enhancement procedure, gaps between the pixels have been bridged using dilation process. One of the advantages of the morphological dilation process over the low pass filtering method is that the morphological method resulted directly in a binary image. Low pass filtering, on the other hand, started with a binary image and produced a gray scale image, which would require a pass with a thresholding function to convert it back to binary form. Statistical tolerance limits has also been applied to obtain the numerals with a standard height, width, and minimum and maximum aspect ratio. The lower tolerance limit can be obtained by applying the formula, *Mean – K * S* and the upper tolerance limit = *Mean + K * S*, where Mean refers to average height of a particular numeral, K=3.291(sample size > 1000 and confidence limit =99.9%) and S is the standard deviation. From the experiment conducted, the tolerance values obtained for validation of the numerals is tabulated in Table 1. Figure 1 shows the sample data set for training which have been obtained after dilation and tolerance limits procedures.

Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1 Shows The Tolerance Limit Height And Width, Lower And Upper Tolerance Limit Of Aspect Ratio Of The Numeral From The Training Data

Numerals	Lower tolerance limit (Height)	Lower tolerance limit (Width)	Lower tolerance limit (Aspect ratio)	Upper tolerance Limit (Aspect ratio)
0	14	9	0.425835	1.820406
1	16	2	0.320	5.399727
2	13	9	0.364921	1.827389
3	15	6	0.40325	2.144832
4	16	7	0.427453	2.136789
5	8	6	0.186222	2.025031
6	13	5	0.468586	2.311148
7	14	6	0.453196	2.166947
8	16	6	0.436202	2.232329
9	16	5	0.565917	2.394791

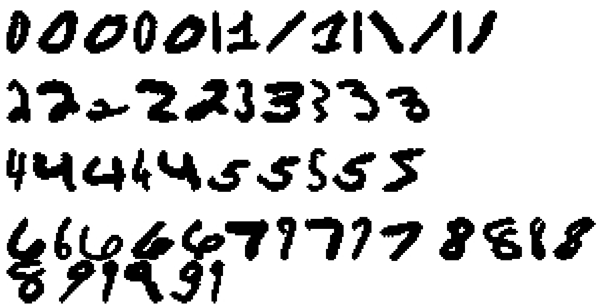


Figure 1 shows sample data set for training which was obtained after dilation and tolerance limits procedure.

In addition to the MNIST database, we have collected Assamese, Devanagari, Manipuri , Malayalam and Oriya handwritten numerals as the primary database. Total number digits collected for Assamese language is 21745, 16200 digits for training set and 5545 digits for testing set. Total digits collected for Devanagari language is 21314, 16200 digits for training set and 5545 digits for testing set. Total digits collected for Manipuri language is 19573, 11700 digits for training set and 7873 digits for testing set. Total digits collected for Malayalam language is 20874, 14400 digits for training set and 6474 digits for testing set. Total digits collected for Oriya language is 21146, 15300 digits for training set and 5846 digits for testing set.

4. REPRESENTATION OF THE HANDWRITTEN DIGITS

Representation plays a vital role in handwritten recognition system. The selection of good representation without affecting the shape of the numerals leads to a better recognition rate. The ratio of Height / Width varies from 0.33 to 5.40 for handwritten numerals due to variance of the numerals of the same class. Based on the aspect ratio midpoint value of a numeral, the number of zones can be obtained along x and y axis. Table 2 shows the class interval, midpoint value and number of zones along x axis and y axis.

Table 2 Shows The Number Of Zones Required Along X Axis And Y Axis Based On Midpoint Values

Serial No	Class Interval (based on aspect ratio)	Mid Point Value	Number of Zones (along y axis)	Number of Zones (along x axis)
1	0.55 – 0.64	0.60	3	5
2	0.65 – 0.74	0.70	7	10
3	0.75 – 0.84	0.80	4	5
4	0.85 – 0.94	0.90	9	10
5	0.95 – 1.04	1.0	4	4
6	1.05 – 1.14	1.10	11	10
7	1.15 – 1.24	1.20	6	5
8	1.25-1.34	1.30	13	10
9	1.35-1.44	1.40	7	5
10	1.45-1.54	1.5	6	4
11	1.55-1.64	1.6	8	5
12	1.65-1.74	1.70	17	10
13	1.75-184	1.80	9	5
14	1.85-1.94	1.90	19	10
15	1.95-2.04	2.0	4	2
16	2.05-2.14	2.10	21	10

5. ZONE BASED FEATURE IDENTIFICATION

The methodology followed for extraction of features based on the above representation scheme plays a vital role in the recognition rate of a numeral. Theoretically,

Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

there are 'M' zones in the set A and 'N' zones in the set B. Then the set of all ordered pairs (a, b) where a ϵ A and b ϵ B, is called the Cartesian product and is denoted by A × B. Since a relation from A to B is precisely a subset of A × B, the set of all relations from A to B is precisely the set of all subsets of A × B. Therefore, the number of relations from A to B is equal to the number of subsets of A × B. Since, number of elements in A is M zones and number of elements in B is N zones, we have A × B = M × N zones. Therefore, A × B has 2^{M*N} number of subsets [25]. Let A = {1, 2, 3…M} and B = {1, 2, 3…N} and let the relation R from A to B be defined as relational density of the zone (feature), i.e., Relational density of the zone = total number of white (black) pixels in the given area / (total number of white pixels + total number of black pixels). Therefore, 2^{M*N}, relational density zone can be obtained from the given M * N zones. [1]

Table 3 Depiction Of Three Row Zones Along Y Axis And Four Column Zones Along X Axis

Z1	Z2	Z3	Z4
Z5	Z6	Z7	Z8
Z9	Z10	Z11	Z12

6. FEATURE SELECTION FOR ZONE BASED REPRESENTATION

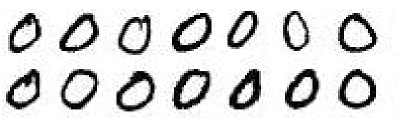
From feature identification procedure, the number of features we could extract is exponential, so, we have narrowed down the number of features based on four primitive descriptors, namely, relational density along x axis, along y axis, along the major diagonal and along the minor diagonal. Hence, the numbers of features become polynomial W. For example, the number of features generated using midpoint value method for the zone size (53, 10) is 83145.

7. PROPOSED METHODOLOGY

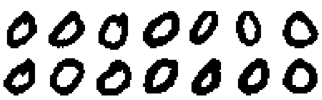
The proposed method uses relative density of the pixels as the main feature in the classification process. Initially, enhancement is carried out for connecting broken numerals using dilation method, the number of zones selection has been done using midpoint class interval values. After selecting the number of zones along x and y axis, standard block size namely (16 *16) or (8*8) of the zone has been assigned and the image is re-sized for feature extraction using interpolation method. This procedure will be done for all similar images of same zone size along both axes. The relative densities of the zone are computed for 1*1, 1*2, 1*3 …….up to M * N. The mean relative density of the pixels for each group (cluster) are found using hierarchical clustering algorithm and finally stored in the feature vector library. To classify the numeral, we have used minimum distance classifier and the nearest feature vector is estimated. The Euclidean distance between the feature vector and the mean feature vector is determined and assigned the numeral class to the nearest mean vector.

Algorithm for Training Numerals
Input: Binary Numeral Image from the database
Output: Store Feature Vector in the Library
Step 1: Pre-process the image to connect the broken digits (dilation)
Step2: Perform labelling connected component algorithm to crop the image.
Step3: Obtain the number of zones along x and y axis based on the midpoint aspect ratio interval
Step 4: Resize the image
Step 5: Compute the relative density of the pixels for $1 \le i \le M$ and $1 \le j \le N$
Step 6: Repeat the steps 1 to 5 to group the images using Hierarchical Clustering Algorithm and store the mean relative density feature vector in the library.
End of the Training Numerals Algorithm.

Algorithm for Testing a Numeral
Input: Isolated Binary Numeral Image
Output: Recognition of the numeral
Step 1: Pre-process the image to connect the broken digits.
Step2: Perform labelling connected component algorithm to crop the image.
Step3: Obtain number of zones along x and y axis based the midpoint aspect ratio interval from the global database
Step 4: Resize the image
Step 5: Compute the relative density of the pixels for $1 \le i \le M$ and $1 \le j \le N$
Step 6: Calculate the Euclidean distance between the input feature vector and the mean feature vectors and assign the class to the nearest mean vector
End of the Testing a Numeral Algorithm.
On the basis of the proposed algorithm, the flow of the system is provided in the following figures and tables.



Sample training data



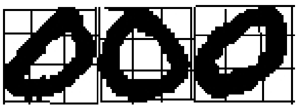
Dilated image

Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Cropped image

The grouped images based on midpoint aspect ratio class
interval



Resized image

The mean relative density of pixels of all the three images

0.6602	0.4844	0.2456	0.5678	0.4326

Figure 2 flow of the system

8. EXPERIMENTAL RESULTS

For the purpose of experimentation, Handwritten Latin, Assamese, Devanagari, Manipuri, Malayalam and Oriya numerals have been used.

In the Experiment I, we have achieved 92.85% as recognition rate for the Latin numerals. In the Experiment II, 99.28% recognition rate has been achieved in Assamese language, In the Experiment III, 98.95% recognition rate has been achieved in Devanagari script, In the Experiment IV, 98.72% recognition rate has been obtained in Manipuri language, In the Experiment V, 99.29% has been obtained for Malayalam Language, In the Experiment VI, 99.48% has been obtained for Oriya language.

Table 4 Sample Confusion Matrix For Latin Numerals (Experiment I)

	0	1	2	3	4	5	6	7	8	9	Accuracy %	#recognized Data	#Test data
0	952	1	0	3	0	0	3	0	1	1	99.06	952	961
1	1	1101	6	2	1	2	5	0	7	0	97.86	1101	1125
2	12	0	904	32	1	6	2	11	13	0	92.15	904	981
3	0	1	12	888	0	29	1	7	33	8	90.70	888	979
4	5	2	6	1	857	3	11	3	6	46	91.17	857	940
5	2	2	4	48	5	706	12	1	24	2	87.59	706	806
6	2	3	3	0	1	13	886	0	1	0	97.46	886	909
7	1	7	7	5	11	1	0	926	4	41	92.32	926	1003
8	17	9	8	27	9	33	7	8	818	16	85.92	818	952
9	4	2	2	16	10	1	0	13	7	910	94.30	910	965
										Mean	92.85		

To prove the eminence of the proposed algorithm a comparative analysis with the popular and best algorithms has been attempted. The table 8 shows the comparison of proposed with existing algorithms for Latin Numerals.

Journal of Theoretical and Applied Information Technology
10th June 2014. Vol. 64 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5 Comparison Of Proposed Method With The Existing Methods For Latin Numerals

References	No. of samples in the Data set	Feature Extraction Method	Classifier	Accuracy %
[1]	9621	Selective Subset of Relative Density Feature Extraction	Minimum Distance Classifier	93.02%
[2]	300	Structural features	Decision tree	96%
[3]	Not available	Contour based on the direction and curvature for a character contour	Basic Multi-layered neural network with three kinds of inputs	98.3%, 98.8%, 99.1%
[4]	2000	Structural features from their skeletons, other makes use of their contours	A tree classifier	84.85%
[6]	500	Not available	Back Propogation Neural network	70%
[8]	11000	Local histogram of the chain codes	Statistical and neural network	96.02%
[11]	10000	concavity measurements, horizontal and vertical projections, polygonization of both the internal and the external numeral boundaries and top, bottom, left and right extrema of the numeral		98.05%
[12]	8000	Aspect invariant, moment order	Neural networks	98.73%
[13]	2600	Shadow coding	Probabilistic Neural Network	97.8%
[14]	10000	Primitive coding and global description	Structural method	99.7%
[15]	17000	Projections and contour histograms	Knn	99.54%
[17]	195000	Foreground features namely, transitions from background and foreground, Background features based on concavity information	HMM	97.9%
[19]	Not available	Statistical and structural features	Neural network	95.7%
[20]	100	Linear discriminant analysis	Improved LDA and Bhattacharyya Distance	96.93%
[22]	18467	Holes, vertical up and down cavities and endpoints, horizontal left, right cavities, horizontal left and right end points	Knn	85.26%
K. N. Saravanan & Dr. R. Anitha	**9621**	**Relative density of pixels**	**Minimum Distance Classifier**	92.8567%

9. CONCLUSION

In this paper we have proposed a relative density feature extraction algorithm for the recognition of single connected component numerals. Minimum distance classifier has been used for classification. The recognition rate of 92.8567%, 99.28%, 98.95%, 98.72%, 99.29%, 99.48% has been achieved for Latin, Assamese, Devanagari, Manipuri, Malayalam and Oriya numerals respectively. The table shows the previous work cited has better performance than the proposed method for Latin Numerals. The reason could be the dataset used and the size of the database. Our future work aims to decrease the number of features and combine other feature set to improve the recognition rate and apply the proposed algorithm to other language numerals and check the robust nature of the algorithm.

10. LIMITATIONS AND FUTURE ENHANCEMENT

The maximum number of features generated using midpoint value method for the zone size (53, 10) is 83145 and (37, 10) is 41665, though it is polynomial, but it occupies more space to store the floating point values. Our future work aims to decrease the number of features which are generated using the proposed algorithm to reduce the space and apply the proposed algorithm to other language numerals.

11. ACKNOWLEDGEMENTS

The authors are grateful to Dr. Nangegowda, Prof. Chandrasekhar, Dr. Fr. Jossy, Dr. Nachamai M. and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Prof Joy Paulose, Department of Computer Science, Christ University, Bangalore, India for their encouragement and constant support throughout this work.

REFERENCES

[1] K.N. SARAVANAN, DR. R. ANITHA, 2014, Selective Subset of Relative Density Feature Extraction Algorithm for Unconstrained Single Connected Handwritten Numeral Recognition, Australian Journal of Basic and Applied Sciences, Pages:429-435.

[2] Kazua KAMATA, Akira WATANABE, Takashi SATOH, and Mikiya SASE,Max Mignotte, 1988, A structural recognition algorithm for handwritten numerals, IEE, pp. 365-368

[3] Keiji YAMADA, Hiroyuki KAMI, Jun TSUKUMO, 1989, Handwritten Numeral Recognition by Multi-Layered Neural Network with Improved Learning Algorithm, Pattern Recognition Research Laboratory

[4] Christine Nadal, Raymond Legault, Ching Y. Suen, 1990, Complementary Algorithms for the Recognition of Totally Unconstrained Handwritten Numerals, IEEE, CH2898-5, pp. 443-449

[5]TUAN A. MAI and CHING Y. SUEN, 1990, A Generalized Knowledge Based System for the Recognition of Unconstrained Handwritten Numerals, IEEE Transactions on Systems, Man, and Cybernetics, vol. 20, No.4, pp.835-848.

[6] J.T. Lin and R.M. Inigo, 1991, Handwritten Zip Code Recognition by Back Propagation Neural Network, IEEE, CH2998-3, pp.731-735

[7] CHING Y. SUEN, CHRISTINE NADAL et., Computer Recognition of Unconstrained Numerals, 1992, Proceedings of the IEEE, Vol. 80. No.7, pp. 1162-1180

[8] Jun Cao, M. Shridhar, F. Kimura, M. Ahmadi, 1992, Statistical and Neural Classification of Handwritten Numerals: A Comparative Study, IEEE, CH0-8186-2915-0, pp.643-646

[9] Heutte L, Moreau J.V., Plessis B, Plagnaud J.L., Lecourtier Y, 1993, Hand printed Numeral Recognition with the Learning Quadratic Discriminant Function, IEEE, CH-0-8186-4960-7, pp.167-170

[10] Yi Lu, Steven Schlosser, Michael Janeczko, Fourier descriptors and handwritten digit recognition, 1993, Machine Vision and Applications, vol. 6, pp. 25-34

[11] Takahiko KAWATANI, 1993, Hand printed Numeral Recognition with the Learning Quadratic Discriminant Function, IEEE

[12] Feng Pan, Mike Keane, 1994, A New Set of Moment Invariants for Handwritten Numeral Recognition, IEEE, CH-0-8186-6950, pp.154-158

[13] M.H. Shirali-Shahreza Karim Faez, Alireza Khotanzad, 1995, Recognition of Handwritten Persian/Arabic Numerals by Shadow Coding and an Edited Probabilistic Neural Network, IEEE, CH-0-8186-7310-9, pp.436-439

[14] Jianming Hu and Hong Yan, 1996, Structural Decomposition and Description of Printed and Handwritten Characters, Proceedings of ICPR, pp.230-234

[15] Thien M. Ha and Horst Bunke, 1997, Off-Line, Handwritten Numeral Recognition by Perturbation Method IEEE Transactions on pattern analysis and machine intelligence Vol. 19, No.5, pp. 535-539

[16] XUEFANG ZHU, C. GEOGER, 1998, An Off-line System for Recognition of Free Written Zipcodes, Proceedings of ICSP, pp.1253-1256

[17] Alceu de S. Britto Jr, Robert Sabouring, Flavio Bortolozzi and Ching Y. Suen, 2003, Complementary Features Combined in an HMM-based System to Recognize Handwritten Digits, Proceedings of the 12th International Conference on Image Analysis and Processing, ICIAP

[18] B.V. Dhandra, V.S. Malemath, Mallikarjun, Ravindra Hegadi, 2006, Multi font Numeral Recognition without Thinning based on Directional Density of Pixels, ICDIM, pp.157-160

[19] Binu P. Chacko, Babu Anto P, 2007, Comparison of Statistical and Structural Features for Handwritten Numeral Recognition, ICCIMA, pp. 296-300

[20] Ying Wen, Pengfei Shi, 2008, An Approach to Numeral Recognition based on Improved LDA and Bhattacharyya Distance, ISCCSP, vol. 12-14, pp.309-311

[21] S.V. Rajashekararadhya, P. Vanaja Ranjan, 2009, Support Vector Machine based Handwritten Numeral Recognition of Kannada Script, IACC, vol. 6, No.7, pp.381-386

[22] S. Impedovo, R. Modugno, G. Pirlo, 2010, Membership Functions for Zoning based Recognition of Handwritten Digits, International Conference on Pattern Recognition, pp.1876-1879

[23] Mahesh Jangid, Renu Dhir, Rajneesh Rani, Kartar Singh, SVM Classifier for Recognition of Handwritten Devanagari Numeral, ICIIP

[24] G. Pirlo, D. Impedovo, 2012, Voronoi based Zoning Design by Multi-Objective Genetic Optimization, IEEE

[25] Md. Musfiqur Rahman Sazal et al, 2013, Bangla Handwritten Character Recognition Using Deep Belief Network, EICT

[26] Ravindra S. Hegadi, Parshuram M. Kamble, 2014, Recognition of Marathi Handwritten Numerals Using Multi-layer Feed Forward Neural Network, WCCCT, pp 21-24.

[27]Dr. D.S. Chandrasekharaiah, 2004, Mathematics: DISCRETE MATHEMATICAL STRUCTURES, Second Edition, Prism Books Private Limited.