Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Zone-Based Hybrid Feature Extraction Algorithm for Handwritten Numeral Recognition of Four Indian Scripts

*S.V.Rajashekararadhya
Research Scholar
Department of Electrical Engineering, CEG
Anna University, Chennai, India
*e-mail: svr_aradhya@yahoo.co.in

Vanaja Ranjan P
Asst.Professor
Department of Electrical Engineering, CEG Anna University, Chennai, India
e-mail: vanaja@annauniv.edu

***Abstract*—India is a multi-lingual and multi-script country, where eighteen official scripts are accepted and there are over hundred regional languages. In this paper we propose a zone-based hybrid feature extraction system. The character centroid is computed and the image (character/numeral) is further divided into n equal zones. An average angle from the character centroid to the pixels present in the zone, is computed (one feature). Similarly, the zone centroid is also computed (two features). The average angle from the zone centroid to the pixels present in the zone is computed (one feature). This procedure is sequentially repeated for all the zones/grids/boxes present in the numeral image. There could be some zones that are empty; then, the value of that particular zone image in the feature vector is zero. Finally, 4*n such features are extracted. The nearest neighbor and support vector machine classifiers are used for subsequent classification and recognition purposes. We obtained 97.85 %, 96.8 %, 95.1% and 95 % recognition rates for Kannada, Telugu, Tamil and Malayalam numerals respectively, using support vector machine.**

***Keywords*—Handwriiten character recognition, image processing, Indian scripts, Feature extraction, Nearest neighbor classifier, Support vector machine**

I. INTRODUCTION

Handwritten character recognition (HCR) is an important area in image processing and pattern recognition fields. HCR has received extensive attention in academic and production fields. The recognition system can be either on-line or off-line. In on-line handwriting recognition, words are generally written on a pressure sensitive surface (digital tablet PCs) from which real time information, such as the order of the stroke made by the writer is obtained and preserved. This is significantly different to off-line handwriting recognition where no dynamic information is available [1]. Off-line handwriting recognition is the process of finding letters and words that are present in the digital image of the handwritten text. It is the subfield of optical character recognition(OCR). Several methods of recognition of English, Latin, Arabic, Chinese scripts are excellently reviewed in [1, 2, 3, 4].

There are five major stages in the HCR problem: Image preprocessing, segmentation, feature extraction, training and recognition and post processing. Research in HCR is popular for various practical application potential such as, reading aids for the blind, bank cheques, vehicle number plates, automatic pin code reading to sort postal mail. There is a lot of demand on Indian scripts character recognition and a review of the OCR work done on Indian languages is excellently reviewed in [5]. In [6] a survey on the feature extraction methods for character recognition is reviewed. The feature extraction method includes Template matching, Deformable templates, Unitary Image transforms, Graph description, Projection Histograms, Contour profiles, Zoning, Geometric moment invariants, Zernike Moments, Spline curve approximation, Fourier descriptors , Gradient feature, Gabor feature etc.

India is a multi-lingual and multi-script country, comprising eighteen official languages, namely, Assamese, Bangla, English, Gujarati, Hindi, Kankanai, Kannada, Kashmiri, Malayalam, Marathi, Nepali, Oriya, Punjabi, Rajasthani, Sanskrit, Tamil, Telugu and Urdu. Recognition of handwritten Indian scripts is difficult because of the presence of numerals, vowels, consonants, vowel modifiers and compound characters.

We will now briefly review the few important works done towards HCR with reference to the Indian language scripts. In [7] the grid based feature extraction method is used to recognize the handwritten Bangla numerals using a multifier classifier. An accuracy of 97.23% is reported. The recognition of conjunctive Bangla Characters by the artificial neural network is reported in [8]. Recognition of Devanagari characters using gradient features and fuzzy-neural network is reported in [9] and [10] respectively. We found curvature feature for recognizing Oriya characters in [11].

In [12] the zone/grid based feature extraction for handwritten Hindi numerals is reported. For extracting the features, each character image is divided into 24 zones. By considering the bottom left corner of the image as an absolute reference, the average vector distance for the pixels present in the zone is computed. Character recognition for Telugu scripts using multi-resolution analysis and associative memory is reported in [13]. The Fuzzy technique and neural network based off-line Tamil character recognition are found in [14].

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

In [15] the author has described a scheme for extracting features from the gray scale images of the handwritten characters on their state-space map with eight directional space variations. In [16] for feature computation, the bounding box of a numeral image is segmented into blocks and the directional features are computed in each of the blocks. These blocks are then down-sampled by a Gaussian filter and the features obtained from the down-sampled blocks are fed to a modified quadratic classifier for recognition. Off-line HCR for numeral recognition is also found in [17-21].

Recognition of isolated handwritten Kannada numerals based on the image fusion method is found in [17]. In [18] Zone and Distance metric based feature extraction method is used. The character centroid is computed and the image is further divided into n equal zones. The average distance from the character centroid to pixels present in the zone is computed. This procedure is repeated for all the zones present in the numeral image. Finally n such features are extracted for classification and recognition.

In [19] the zone and vertical projection distance metric is used. The preprocessed numeral image (50x50) is fed to the feature extraction module. The image is divided into 25 zones (each zone size is 10x10). The average pixel distance of each column present in the zone is computed vertically. In [20] image is further divided in to n equal zones. The average distance from the zone centroid to pixels present in the zone is computed.

From the above literature survey it is clear that not much work has been done on the south-Indian scripts. This motivated us to work on south-Indian scripts. The selection of the feature extraction method is also a very important factor for achieving efficient character recognition. In this paper we, propose a simple and efficient zone-based hybrid feature extraction algorithm. The Nearest Neighbor Classifier (NNC) and Support Vector Machine (SVM) classifiers are used for the recognition and classification of a numeral image. We have tested our method for Kannada, Telugu, Tamil and Malayalam numerals and also experimentation is carried out on MNIST digit database and ISI Bangla digit database.

The rest of the paper is organized into six sections. In Section II we shall briefly explain the overview of Kannada, Telugu, Tamil and Malayalam scripts. In Section III we shall briefly explain the data collection and preprocessing. In Section IV we shall discuss the proposed feature extraction method. Section V describes the numeral classification and recognition using different classifiers, experimental results and comparative study, and finally, the conclusion is given in Section VI.

II. Overview of Kannada, Telugu, Tamil and Malayalam scripts

In this section, we shall explain the properties of four popular south-Indian scripts. Most of the Indian scripts have originated from the Brahmi script through various transformations. The writing style of Indian scripts considered in this paper is from left to right, and the concept of the upper/lower case is not applicable to these scripts.

Kannada is one of the major Dravidian languages of South India and one of the earliest languages evidenced epigraphically in India, and spoken by about 50 million people in the Indian states of Karnataka, Tamil Nadu, Andhra Pradesh and Maharasthra. The script has 49 characters in its alphasyllabary and is phonetic. The characters are classified into three categories: swaras(vowels), vyanjans(consonants) and yogavaahas (part vowel, part consonants). The script also includes 10 different Kannada numerals of the decimal number system.

Telugu is a Dravidian language and it is the third most popular script in India. It is the official language of the south Indian state, Andhra Pradesh and also spoken by neighboring states. The Telugu script is closely related to the Kannada script. Telugu is a syllabic language. Similar to most languages of India, each symbol in the Telugu script represents a complete syllable. Officially, there are 10 numerals, 18 vowels, 36 consonants, and three dual symbols.

Tamil is a Dravidian language and one of the oldest languages in the world. It is the official language of the Indian state of Tamil Nadu; it also has official status in Sri Lanka, Malaysia and Singapore. The Tamil script has 10 numerals, 12 vowels, 18 consonants and five grantha letters. The script, however, is syllabic and not alphabetic. The complete script, therefore, consists of 31 letters in their independent form, and an additional 216 combining letters representing every possible combination of a vowel and a consonant.

Malayalam is a Dravidian language and it is the eighth most popular script in India, and spoken by about 30 million people in the Indian state of Kerala. Both the language and the writing system are closely related to Tamil. However, Malayalam has its own script. The script has 16 vowels, 37 consonants and 10 numerals.

The challenging part of the Indian handwritten numeral/character recognition is the distinction between the similar-shaped components. A very small variation between two characters or numerals leads to recognition complexity and a certain degree of recognition accuracy. The style of writing the characters is highly different and they come in various sizes and shapes. The same numeral may take different shapes, and conversely, two or more different numerals of a script may take similar shape.

III. Data collction and preprocessing

Data collection for the experiment has been done from different individuals. Currently we are developing data set for Kannada, Telugu, Tamil and Malayalam numerals.

Earlier we had collected 2000 Kannada numeral samples from 200 different writers [21]. Writers were provided with a plain A4 sheet and each writer asked to write Kannada numerals from 0 to 9 at one time. Recently, we have again collected 2000 Kannada numerals by 40 different writers. In this paper the data set size of 4000 Kannada numerals is used. The database is totally unconstrained and has been created for validating the recognition system. Similarly, we have collected 2000 Telugu numeral samples from 200 different writers, 2000 Tamil numeral samples from 200 writers and 1500 Malayalam numeral samples from 150 different writers. The collected

documents are scanned using the HP-scan jet 5400c at 300dpi, which is usually a low noise and good quality image. The digitized images are stored as binary images in the BMP format. A sample of Kannada, Telugu, Tamil and Malayalam handwritten numerals from the data set are shown from figures 1 to 4 respectively.

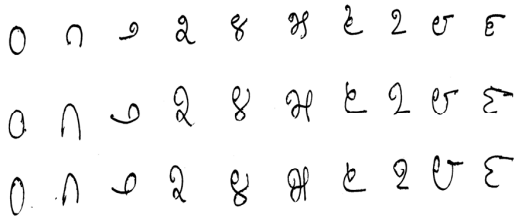

Figure 1. Sample of handwritten Kannada numerals from 0 to 9

Figure 2. Sample of handwritten Telugu numerals from 0 to 9

Figure 3. Sample of handwritten Tamil numerals from 0 to 9

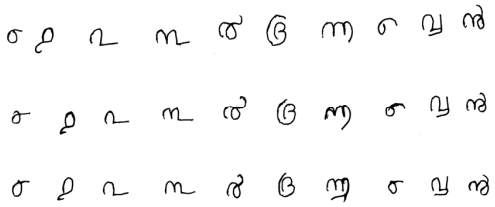

Figure 4. Sample of handwritten Malayalam numerals from 0 to 9

Preprocessing includes the steps that are necessary to bring the input data into an acceptable form for feature extraction. The raw data, depending on the data acquisition type, is subjected to a number of preliminary processing stages. The preprocessing stage involves noise reduction, slant correction, size normalization and thinning. Among these, size normalization and thinning are very important. Normalization is required as the size of the numeral varies from person to person and even with the same person from time to time. The input numeral image is normalized to size 50x50 after finding the bounding box of each handwritten numeral image.

Thinning provides a tremendous reduction in data size; it extracts the shape information of the characters. It can be considered as the conversion of off-line handwriting to almost on-line data. Thinning is the process of reducing the thickness of each line of pattern to just a single pixel. In this research work, we have used the morphology based thinning algorithm for better symbol representation. The detailed information about the thinning algorithm is available in [22].

Thus, the reduced pattern is known as the skeleton and is close to the medial axis, which preserves the topology of the image. Figure 5 shows the steps involved in our method as far as preprocessing is considered.

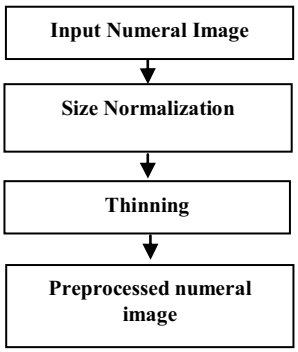

Figure 5. Preprocessing of the input numeral image

IV. PROPOSED FEATURE EXTRACTION METHOD

For extracting the feature, the zone-based hybrid approach is proposed. The most important aspect of the handwriting recognition scheme is the selection of a good feature set, which is reasonably invariant with respect to shape variations caused by various writing styles. The major advantage of this approach stems from its robustness to small variations ease of implementation and good recognition rate. The zone-based feature extraction method gives good results even when certain preprocessing steps like filtering; smoothing and slant removing are not considered. In this section, we shall explain the concept of the feature extraction method used for extracting features for efficient classification and recognition. The following paragraph explains in detail, the proposed feature extraction methodology.

The character centroid is computed and the image (50x50) is further divided into fifty equal zones as shown in figure. 6. The average angle from the character centroid to the pixels present in the zone is computed (one feature) (ICZA). Similarly zone centroid (ZC) is computed (two features). The average angle from the zone centroid to the pixels present in the zone is computed (one feature) (ZCZA). This procedure is sequentially repeated for all the zones/grids/boxes present in the numeral image. There could be some zones that are empty, and then the value of that particular zone image in the feature vector is zero. Finally 200 such features are used for feature extraction. For classification and recognition, NNC and SVM

classifiers are used. Proposed algorithm provides the hybrid feature extraction system (ICZA-ZCZA-ZC).

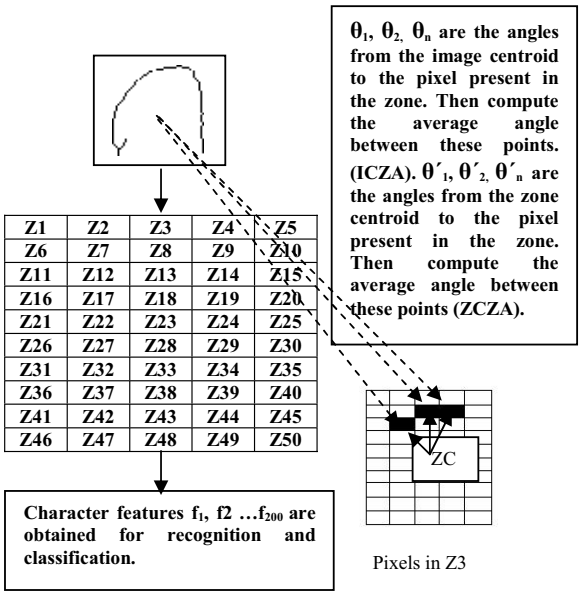

Figure 6. Procedure for extracting features from the numeral image

V. CLASSIFIERS, EXPERIMENTAL RESULTS AND COMPARATIVE STUDY

A. Nearest neighbor classifier for classification and recognition

For large-scale pattern matching, a long-employed approach is the NNC. The training phase of the algorithm consists only of storing the feature vectors of the training samples. In the actual classification phase, the same features as before are computed for the test samples. Distances from the new vector to all the stored vectors are computed. Then, classification and recognition is achieved on the basis of similarity measurement.

B. Support vector machine for classification and recognition

The support vector machine is a new classifier that is extensively used in many pattern recognition applications. The SVM uses the principle of structural risk minimization by minimizing Vapnik Chervonenkis (VC) dimensions [25, 26]. Regarding the pattern classification problem, the SVM demonstrates a very good generalization performance in empirical applications.

SVM is binary classifier that separates linearly any two classes by finding a hyper plane of maximum margin between the two classes. The margin means the minimal distance from the separating hyper plane to the closest data points. The SVM learning machine searches for an optimal separating hyper plane, where the margin is maximal. The outcome of the SVM is based only on the data points that are at the margin and are called support vectors. There are two approaches to extend the SVM for multi-class classification. The first one is one against one (ONO) and the other is one against all (ONA). We have used the ONA approach where N SVM classifiers are performed to separate one of N mutually exclusive classes from all other classes.

A kernel is utilized to map the input data to a higher dimensional feature space, so that the problem becomes linearly separable. The kernel plays a very important role. The Gaussian kernel performs better compared to the linear kernel, polynomial kernel etc. We have used the Gaussian kernel.

C. Experimental results

For recognition and classification purposes NNC and SVM classifiers are used. Table I gives the results of confusion matrix for Kannada numerals for different 2000 training and 2000 testing samples using SVM classifier. Table II gives the results of confusion matrix for Telugu numerals for different 1000 training and 1000 testing samples using SVM classifier.

TABLE I. CONFUSION MATRIX FOR KANNADA HANDWRITTEN NUMERALS USING SVM CLASSIFIER

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**99.5**			0.5						
1		**100**								
2			**100**							
3			1	**94**	1	0.5	0.5	3		
4					**99**				0.5	0.5
5		0.5		2	1	**96**	0.5			
6							**96**	3	1	
7		0.5		1.5				**98**		
8	0.5		0.5	1					**97.5**	0.5
9							1	0.5		**98.5**
Average Recognition Rate for Kannada numerals = 97.85 %										

TABLE II. CONFUSION MATRIX FOR TELUGU HANDWRITTEN NUMERALS USING SVM CLASSIFIER

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**94**	4	1				1			
1	4	**96**								
2	3		**97**							
3			1	**98**		1				
4					**98**	2				
5			1		2	**97**				
6							**100**			
7				3	1		4	**92**		
8									**98**	2
9							2			**98**
Average Recognition Rate for Telugu numerals = 96.8 %										

TABLE III. CONFUSION MATRIX FOR TAMIL HANDWRITTEN NUMERALS USING SVM CLASSIFIER

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**97**				1	1		1		
1	1	**96**	1							2
2			**97**	3						
3		1	1	**92**	5		1			
4					**99**				1	
5	1				2	**96**		1		
6					2	1	**92**		1	4
7		1			1			**97**	1	
8		1						2	**97**	
9		1		1			8		2	**88**
Average Recognition Rate for Tamil numerals = 95.1 %										

Table III gives the results of confusion matrix for Tamil numerals for different 1000 training and 1000 testing samples

using SVM classifier. Table IV gives the results of confusion matrix for Malayalam numerals for different 1000 training and 500 testing samples using SVM classifier. The most confusing Tamil numeral pair is 6 and 9. Due to similarity among these two numerals, the over-all recognition rate is reduced. The SVM classifier gives better results compare to the NNC classifier. Similarly, in Malayalam numerals also the most confusing numeral is 4 and is misclassified as 9.

TABLE IV. CONFUSION MATRIX FOR MALAYALAM HANDWRITTEN NUMERASL USING SVM CLASSIFIER

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**94**							6		
1		**98**						2		
2			**100**							
3			2	**96**						2
4		4			**78**			2		16
5						**98**		2		
6					2	2	**94**		2	
7	2		2			2	2	**92**		
8									**100**	
9										**100**
Average Recognition Rate for Malayalam numerals = 95 %										

D. Comapative study

We compared our results with some of the published work on off-line handwritten numerals of Indian scripts. In [12] for Hindi numerals, the zone-based feature extraction algorithm is used and they obtained 95% recognition accuracy. We have implemented the same feature extraction algorithm [12] and we tested it with 4000 Kannada numeral samples. We obtained a recognition rate of 96.4% using the NNC classifier. We obtained 97.70 % recognition accuracy for the proposed algorithm. Table V provides the comparative results for a common data set. Similarly table VI, table VII and table VIII provides comparative results for Telugu, Tamil and Malayalam numerals respectively.

TABLE V. COMPARATIVE RESULTS FOR KANNADA NUMERALS. CV REFERS TO CROSS VALIDATION. BPNN REFERS TO FEED FORWARD BACK PROPAGATION NEURAL NETWORK CLASSIFIER.

Kannada numerals data set size = 4000		
Feature extraction method	***Classifier***	***Recognition rate (%)***
Zoning [12] this paper	NNC	96.4
Zoning [12]this paper CV	NNC	96.05
ZC-ICZ [24] Dataset size 2000	NNC	95.3
ZC-ICZ [24]This paper	NNC	97.35
ZC [24]This paper	NNC	96.8
ZC-ICZA [27]	NNC	97.25
ZC-ICZA [27] CV	NNC	96.78
ZC-ICZA [27]	BPNN	94.75
ZC-ICZA [27]	SVM	97.3
ICZA-ZCZA-ZC Proposed	NNC	**97.7**
ICZA-ZCZA-ZC Proposed	SVM	**97.85**
ICZA-ZCZA-ZC (CV) Proposed	NNC	**96.9**
ICZA-ZCZA-ZC (CV) Proposed	SVM	**97.45**

Also we have used the two fold cross validation scheme for recognition result evaluation. Here the database of the Kannada script is divided into 2 subsets and testing is done on each subset using the rest of the subset for learning. The recognition rates of all the 2 test subsets of the data set are averaged to get the recognition result. Similarly, cross validation is also performed for the Telugu and Tamil numerals data set.

TABLE VI. COMPARATIVE RESULTS FOR TALUGU NUMERALS. CV REFERS TO CROSS VALIDATION. BPNN REFERS TO FEEDFORWARD BACK PROPAGATION NEURAL NETWORK CLASSIFIER

Telugu numerals data set size = 2000		
Feature extraction method	***Classifier***	***Recognition rate (%)***
Zoning [12] this paper	NNC	91.5
Zoning [12] CV This paper	NNC	92.7
ZC [24]	NNC	92.4
ZC-ICZ [24]	NNC	93.8
ZC-ICZA [27]	NNC	94.4
ZC-ICZA [27] CV	NNC	94.6
ZC-ICZA [27]	BPNN	94
ZC-ICZA [27]	SVM	96.2
ICZA-ZCZA-ZC Proposed	NNC	**94.5**
ICZA-ZCZA-ZC Proposed	SVM	**96.8**
ICZA-ZCZA-ZC(CV) Proposed	NNC	**94.9**
ICZA-ZCZA-ZC (CV) Proposed	SVM	**96.85**

TABLE VII. COMPARATIVE RESULTS FOR TAMIL NUMERALS. CV REFERS TO CROSS VALIDATION. BPNN REFERS TO FEEDFORWARD BACK PROPAGATION NEURAL NETWORK CLASSIFIER

Tamil numerals data set size = 2000		
Feature extraction method	***Classifier***	***Recognition rate (%)***
Zoning [12] this paper	NNC	90.6
Zoning [12] **CV** - This paper	NNC	90.6
ZC [24] -This paper	NNC	93.5
ZC [24] CV -This paper	NNC	93.95
ZC-ICZA [27]	NNC	91.3
ZC-ICZA [27] CV	NNC	91
ZC-ICZA [27]	BPNN	90
ZC-ICZA [27]	SVM	93.5
ICZA-ZCZA-ZC -Proposed	NNC	**92.5**
ICZA-ZCZA-ZC - Proposed	SVM	**95.1**
ICZA-ZCZA-ZC(CV) Proposed	NNC	**92.5**
ICZA-ZCZA-ZC(CV) Proposed	SVM	**94.45**

TABLE VIII. COMPARATIVE RESULTS FOR MALAYALAM NUMERALS.BPNN REFERS TO FEEDFORWARD BACK PROPAGATION NEURAL NETWORK CLASSIFIER

Malayalam numerals data set size =1500		
Feature extraction method	***Classifier***	***Recognition rate (%)***
Zoning [12] this paper	NNC	93.2
ZC [24]	NNC	90.8
ZC-ICZ [24]	NNC	90.2
ZC-ICZA [27]	NNC	93.6
ZC-ICZA [27]	BPNN	90.2
ZC-ICZA [27]	SVM	93.2
ICZA-ZCZA-ZC Proposed	NNC	**94.6**
ICZA-ZCZA-ZC Proposed	SVM	**95**

E. Experimental result on ISI Bangla numeral database

In this section, we experimentally evaluate the performance of proposed method on well-known ISI Bangla numerals database [28]. The ISI Bangla numeral database has 19,392 training samples and 4000 test samples. In

preprocessing stage, we performed image size normalization (50x50) and thinning of ISI Bangla handwritten digits. We have considered 12000 training samples and 3000 testing samples for our experimental analysis. For the proposed feature extraction method, we achieved **94.2 %** recognition rate using SVM classifier.

F. Experimental result on MNIST database

In this section, we experimentally evaluate the performance of proposed method on well-known MNIST database (http://yann.lecun.com/exdb/mnist) of handwritten digits. MNIST database consists of 60,000 training samples and 10,000 testing samples. All the digits have been size normalized and centered in 28–by–28 box. In preprocessing stage, we performed image size normalization (50x50) and thinning of MNIST handwritten digits. We have considered 5000 training samples and 1000 testing samples for our experimental analysis. For the proposed feature extraction method, we achieved **95.5 %** recognition rate using SVM classifier.

VI. CONCLUSION

In this paper we have proposed the hybrid type zone-based feature extraction algorithm for the recognition of four popular Indian numeral scripts. The nearest neighbor and support vector machine classifiers are used for subsequent classification and recognition. We have obtained a maximum recognition rate of 97.85% for Kannada numerals using support vector machine.

Our future work aims to develop new zone-based feature extractions algorithms, which produce efficient results. In addition, we plan to extend our work to larger a data set. Also the effective implementation of hybrid classifier system is one of our future research directions.

REFERENCES

[1] R. Plamondon and S. N. Srihari, "On-line and off- line handwritten character recognition: A comprehensive survey", IEEE. Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63-84, 2000.

[2] Nafiz. Arica and Fatos T. Yarman-Vural, "An Overview of character recognition focused on off-line handwriting", IEEE Transactions on System.Man.Cybernetics-Part C: Applications and Reviews, vol. 31, no. 2, pp. 216-233, 2001.

[3] Liana M. Lorigo and Venu Govindaraju, "Offline Arabic handwriting recognition: A survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 5, pp. 712-724, 2006.

[4] G. Nagy, "Chinese character recognition, A twenty five years retrospective", Proceedings of ICPR, pp. 109-114, 1988.

[5] U. Pal, B. B. Chaudhuri, "Indian Script Character recognition: A survey", Pattern Recognition, vol. 37, pp. 1887-1899, 2004.

[6] Anil.K.Jain and Torfinn Taxt, "Feature extraction methods for character recognition-A Survey", Pattern Recognition, vol. 29, no. 4, pp. 641-662., 1996,

[7] A. Majumdar and B.B. Chaudhuri, "Printed and handwritten Bangla numeral recognition using multiple classifier outputs", Proceedings of the first IEEE ICSIP06, Vol. 1, pp. 190-195, 2006.

[8] Abdur Rahim, Shuvabranta Saha, Mahfuzur and Abdus sattar, "Recognition of conjunctive Bangla characters by artificial neural network", International Conference on Information and Communication Technology, pp. 96-99, 2007.

[9] U. Pal, N.Sharma, T. Wakabayashi and F. Kimura, "Off line handwritten character recognition of Devanagiri Scripts'', Ninth International Conference on Document Analysis as Recognition (ICDAR 2007), pp.496-500, 2007.

[10] P.M. Patil, T.R. Sontakke, "Rotation scale and translation invariant handwritten Devanagiri numeral character recognition using fuzzy neural network", Elsevier, Pattern Recognition, vol. 40, pp. 2110-2117. 2007.

[11] U. Pal, T. Wakabayashi and F. Kimura, " A system for off-line Oriya handwritten character recognition using curvature feature ", 10th International Conference on Information Technology, IEEE, pp.227-229, 2007.

[12] M. Hanmandlu, J. Grover, V. K. Madasu and S. Vasikarla, "Input fuzzy for the recognition of handwritten Hindi numeral:", International Conference on Informational Technology, vol. 2, pp. 208-213., 2007.

[13] Arun. K. Pujari, C. Dhanunjaya, Naidu, M. Sreenivasa, Rao and B.C. Jinaga, " An Intelligent character recognizer for Telugu scripts using multiresolution analysis and associative memory", Elsevier, Image vision Computing, pp.1221-1227, 2004.

[14] R.M.Suresh, S.Arumugam, "Fuzzy technique based recognition of handwritten characters", Elsevier, Image Vision Computing, Vol. 25, pp. 230-239, 2007

[15] Lajish. V. L, " Handwritten character using gray- scale based state-space parameters and class modular neural network ", *IEEE* International Conference on signal processing, Comunication and Networking, pp. 374-379, 2008.

[16] U. Pal, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts", Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, pp.749-753, 2007.

[17] G.G. Rajaput and Mallikarjun Hangarge, "Recognition of isolated handwritten Kannada numerals based on image fusion method: ", PReMI07, LNCS.4815, pp.153-160, 2007.

[18] S.V. Rajashekararadhya and P. Vanaja Ranjan, "Isolated handwritten Kannada digit recognition: A novel approach", Proceedings of the International Conference on Cognition and Recognition", pp.134-140, 2008.

[19] S.V. Rajashekararadhya, P. Vanaja Ranjan and V.N. Manjunath Aradhya, "Isolated handwritten Kannada and Tamil numeral recognition: A novel approach", First International Conference on Emerging Trends in Engineering and Technology ICETET 08, pp.1192-1195, 2008.

[20] S.V. Rajashekararadhya, and P. Vanaja Ranjan ,"Handwritten numeral recognition of three popular South Indian scripts: A novel approach:", Proceedings of the second International Conference on information processing ICIP, pp.162-167, 2008.

[21] S.V. Rajashekararadhya, and P. Vanaja Ranjan ,"Neural network based handwritten numeral recognition of Kannada and Telugu scripts", TENCON 2008 Hyderabad – pp.1-5, 2008.

[22] Rafael C. Gonzalez, Richard E. woods and Steven L. Eddins, Digital Image Processing using MATLAB, Pearson Education, Dorling Kindersley, South Asia, 2004.

[23] S.V. Rajashekararadhya, and P. Vanaja Ranjan ,"Efficient handwritten numeral recognition of Kannada and Telugu scripts", International conference on sensors, security, software and intelligent systems (ISSSIS09), pp.IP24-IP28, 2009.

[24] S.V. Rajashekararadhya, and P. Vanaja Ranjan ,"Efficient zone based feature extraction algorithms for handwritten numeral recognition of Indian scripts:", International journal of information processing IJIP, Volume 2 No 4, pp.15-28 2008

[25] V.N. Vapnik, Statistical Learning Theory. John Wiley and sons, 1998.

[26] V.N. Vapnik, The nature of Statistical Learning Theory. Springer, New York, 2nd edition, 1999

[27] S.V. Rajashekararadhya, and P. Vanaja Ranjan,"A novel zone based feature extraction algorithm for handwritten numeral recognition of four Indian scripts:", Digital technology journal, DTJ, Technical University of Ostrava, Volume 2 pp. 41-51, 2009

[28] U. Bhattacharya, and B.B. Chaudhuri,"Handwritten Numeral databases of Indian scripts and multistage recognition of mixed numerals:", IEEE Transaction on Pattern Recognition and Machine Intelligence, Volume 31 No 3, pp. 444-457, 2009