ISSN No. 0976-5697

Volume 3, No. 7, Nov-Dec 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Malayalam Character Recognition System for Camera Enabled Mobile Devices

Jomy John*, Kannan Balakrishnan, Pramod K. V.
Department of Computer Applications
Cochin University of Science and Technology
Kochi, India
jomyeldos@gmail.com*, mullayilkannan@gmail.com, pramodkv4@gmail.com

Abstract: In this paper, we propose a character recognition system for mobile camera phones. The system has modules, namely, image pre-processing, image segmentation, feature extraction, and classification. Initially, contrast stretching is performed on the character image captured through mobile camera and it is segmented and stored in the database as a gray scale image. Then, features using different wavelet filters are extracted from gray scale image of characters. For classification, discriminative power of support vector machine is used. The results shows SVM with RBF kernel yield the best performance with 93.12% of accuracy in recognizing isolated characters.

Keywords: character recognition, camera-based image, wavelet features, SVM, RBF

I. INTRODUCTION

Character recognition is the process of converting an image representation of a document into digital form. The basic task is to assign each character to its class. The document image can be printed or handwritten. Handwritten data is converted to digital form either by scanning the writings on paper or by writing with a special pen on an electronic surface such as a digitizer combined with a liquid crystal display. The two approaches are termed as off-line and on-line handwriting, respectively. The order of strokes made by the writer is available in the latter but only the completed image is available in the former [1]. Another major usage of character recognition technology is in the application of mobile or handheld devices [2]. Mobile OCR will convert scanned documents from camera or photo album into a digital text, which is very much useful for photos of documents one capture while on travelling.

A few other application include recognizing characters of automobile number plates, reading aid for blind or persons having low vision, extracting information from business card, mobile based language learning, etc. All of these applications ultimately lead to the machine's ability to recognize characters. The advantage of mobile OCR is basically due to its portability, flexibility and ease of use. They can be applied to image documents which are difficult to scan, such as text on buildings, vehicles or other objects moving in a scene [2]. However, the process of recognizing characters captured with mobile cameras is not negligible. Major challenges are due to perspective distortion, uneven illumination and shadows, low quality images etc. In addition to this, most mobile devices have less computing power and insufficient storage space. This paper presents an efficient method for classifying Malayalam characters captured through mobile camera. We focus on the images having non uniform illumination, extraction of relevant wavelet features and finally classification using support vector machine.

Malayalam is one of twenty two scheduled languages of India, with rich literary heritage. It has official language status in the State of Kerala and Union territories of Lakshadweep and Mahe. The language is spoken by around 35 million people and it is ranked eighth in terms of the number of speakers in India. Malayalam script is derived from the Grantha script, an inheritor of olden Brahmi script. Most of Malayalam alphabets have circular shapes. It is syllabic in nature and alphabets are classified into vowels and consonants [3]. The script consists of 15 vowels and 36 consonants. In addition to this basic set, the script contains vowel modifiers, half consonants and a large number of conjunct characters. The script also contains 10 numerals, but it is seldom in use. Instead Arabic numerals are used in practice.

The most important issues of developing a handwriting recognition system are the choice of a discriminative feature set and designing a classifier. In this paper, we adapt features from discrete wavelet transform as they provide powerful tool for multi-resolution analysis of images. The application of wavelet transform can be seen in a wide range of problems. A method for localization of licence plates using discrete wavelet transform was proposed by Yuh-Rau Wang et al. [4]. Wunsch and Laine [5] used wavelet features extracted from contour of the handwritten characters for classification using neural networks. Lee et al. [6] used wavelet features extracted from handwritten numerals are classified it using multilayer cluster neural network. Chen et al. [7] developed multi-wavelet descriptor for contour of handwritten numerals using neural network.

Popular classifiers include multi layer perceptron (MLP), learning vector quantization (LVQ), extreme learning machine (ELM), modified quadratic discriminant function (MQDF), radial basis function (RBF) and support vector machine (SVM). We have chosen SVM as it is an effective discriminative classifier with good generalizations and convergence property. Liu et al. [8] have evaluated the

© 2010, IJARCS All Rights Reserved

performance of several classifiers including SVM on CENPARMI, CEDAR and MNIST handwritten numeral databases. Bellili et al. [9] introduced a hybrid Multilayer Perceptron (MLP) and SVM classifiers for handwritten digit recognition. Handwritten Tamil Character Recognition system using SVM classifiers were proposed by Shanthi et al. [10]. Bhowmik et al. [11] provides SVM based hierarchical classification schemes for recognition of handwritten Bangla characters. To our best knowledge, the use of SVM in offline handwritten Malayalam characters was reported only in our previous paper [12].

Most of the previous works require a high resolution scanned image of characters and only a few number of research works can be traced on mobile OCR. Laine et al. [13] developed a system for capital letters in English after skew correction and segmentation. However, the accuracy obtained is not satisfactory. Koga et al. [14] outlined a Kanji OCR for recognizing machine printed Japanese characters and translating them into English. Bae et al. [15] proposed an embedded optical character system of camera captured document image. In [16], the authors presented a mobile application that integrates a touch-based user interface with intelligent character recognition techniques in which recognition is achieved by employing support vector machine. Hardly any work can be traced in Malayalam mobile character recognition.

The paper is organized as follows: the next section discusses the proposed system architecture; section 3 discusses pre-processing phase which includes image acquisition, enhancement, binarization and segmentation; section 4 explores feature extraction using different families of wavelet filters; section 5 discusses SVM, section 6 evaluates the system with experimental results and section 7 concludes the paper.

II. PROPOSED SYSTEM ARCHITECTURE

The proposed system uses a client server architecture, which allows the user to snap image of handwriting and send it to the server; a server side program invokes character recognition procedure to recognise the characters; server sends back the digital text of handwriting back to the phone, where it is displayed. Fig. 1 displays the overall system architecture. The framework of character recognition system in the server is schematically illustrated in Fig. 2

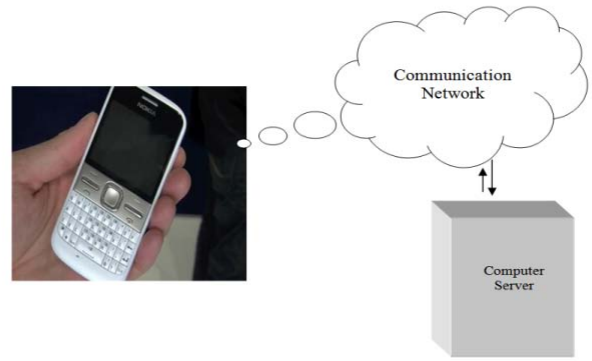

Figure 1. Overall system architecture

III. PRE-PROCESSING PHASE

Pre-processing steps are pre requisite for extracting more meaningful features and it helps improve the recognition accuracy. In the following subsections, steps of the recognition system are described in detail.

A. Image Acquisition and Enhancement:

The image is captured through a mobile camera (Nokia E5-00) keeping it parallel so as to avoid perspective distortion. However, shading and illumination cannot be controlled easily. Low contrast of intensity on image exists due to the illumination variation. To conquer this problem, after converting the color image to gray scale, image enhancement is performed by contrast stretching as in (1), where ‘min’ is the minimum intensity and ‘max’ is the maximum intensity of f(x,y), the input image.

$$g(x,y) = 255*(f(x,y) - \min)/(\max - \min) \tag{1}$$

The obtained image g(x,y) is smoothed to remove noise using a median filter of size 3x3.

B. Segmentation:

Text lines from the binary image are separated and extracted from the $M \times N$ bitmap of the whole image $img(x,y)$, $1 \le x \le M; 1 \le y \le N$ The horizontal projection profile is computed using the following function.

$$H(x) = \sum_{i=1}^{N} img(x,i) \tag{2}$$

From this profile, the peak-valley points are identified and that is used for line separation. The characters in each extracted line are labeled using connected component labeling algorithm. The coordinates of the minimum bounded rectangle containing the character component is identified and the character is segmented from the original gray scale image. The segmented character is stored in gray scale itself in the database.

© 2010, IJARCS All Rights Reserved

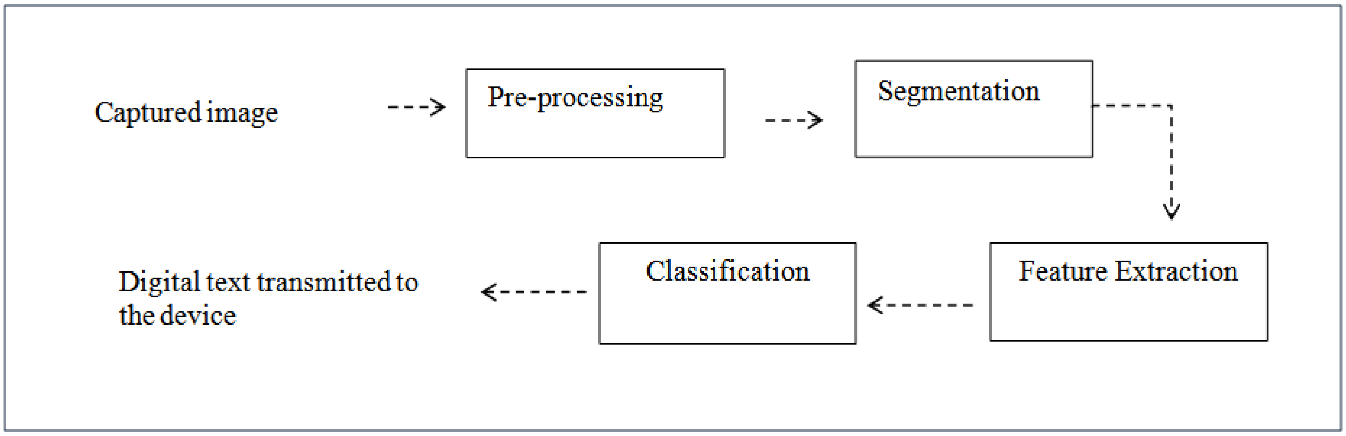

Figure. 2 Block diagram of a character recognition system

IV. FEATURE EXTRACTION USING WAVELET TRANSFORM

A. Wavelet transform:

The space frequency localization and multi resolution analysis capability of a wavelet makes it an efficient tool in analyzing images [17,18]. Wavelet transform decomposes an image into a set of different resolution sub-images, corresponding to the various frequency bands. This results in space frequency localization which is helpful for extracting relevant features.

Discrete wavelet transform can be obtained using the analysis filters for decomposition and the synthesis filters for reconstruction. The scaling function φ(x) and the wavelet function ψ(x) associated with the scaling filter h_φ and the wavelet filter h_ψ are:

$$\varphi(x) = \sum_n h_\varphi(n)\sqrt{2}\varphi(2x-n) \tag{3}$$

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\varphi(2x-n) \tag{4}$$

The multi-resolution technique can be implemented using sub-band decomposition in which the image of a character is decomposed into wavelet coefficients [18]. The rows and columns of the original image is convolved with low pass filter h_φ and high pass filter h_ψ followed by decimation by a factor of two in each direction to generate lower scale components namely low-low(LL), and low-high(LH), high-low(HL) and high-high(HH) sub-images. Three of them, LH, HL and HH correspond to the high resolution wavelet coefficients in the horizontal, vertical and diagonal directions respectively.

LL image is the approximation of the original image and all the four of them contain one-fourth of the original number of samples. As images are very rich in low frequency content, we do analysis further by decomposing low pass filtered version of the image as termed as dyadic partitioning. Fig. 3 explains the decomposition, in which j+1 stands for the starting scale, m and n are row and column directions.

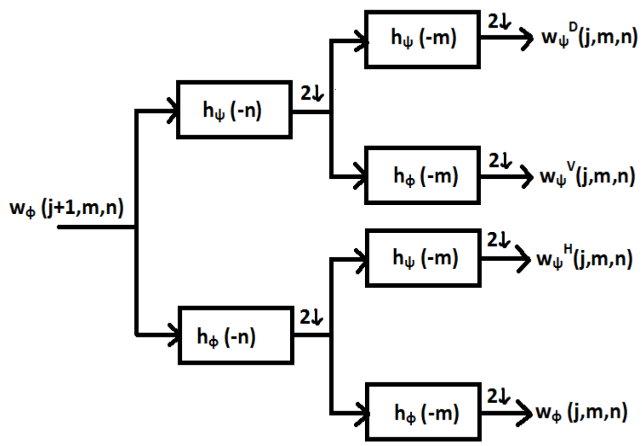

Figure.3 Decomposition using analysis filter bank

B. Feature Extraction:

Feature extraction is crucial for any character recognition system, in which the characters are represented by a set of features. The goal of feature extraction is to find a mapping from the two dimensional image into a smaller one dimensional feature vector $X^T = (x_1, \ldots, x_m)$, that extracts most of the relevant information of the image. The purpose of the feature extractor is to make intra-class variance small by making large inter-class separation. This means that features extracted from samples of same class should be similar, while that of different classes should be dissimilar.

For feature extraction, orthogonal and bi-orthogonal wavelets are used. Orthogonal and compactly supported wavelets include Haar wavelet, Daubechies wavelets, symlets and coiflets. Haar wavelet is a discontinuous compactly supported symmetric orthogonal wavelet. It can be considered as Daubechies-1, with support width of 1. Daubechies wavelets are compactly supported orthogonal wavelets with highest number of vanishing moments for a given support width. Associated scaling filters are minimum-phase filters. Symlets wavelets are compactly supported wavelets with near symmetry and highest number of vanishing moments for a given support width. Associated scaling filters are near linear-

© 2010, IJARCS All Rights Reserved

phase filters. Coiflets wavelets are compactly supported wavelets with highest number of vanishing moments for both φ and ψ for a given support width. Biorthogonal wavelets are compactly supported biorthogonal spline wavelets for which symmetry and exact reconstruction are possible [17].

The original image, which is stored in gray scale is size normalized to 64x64 pixels. The wavelet decomposition is applied to this gray scale image to yield four 32x32 sub images at LL_1, LH_1, HL_1 and HH_1. During the next level decomposition, it yields a 16x16 image [LL_2, LH_2, HL_2 and HH_2] and then an 8x8 image [LL_3, LH_3, HL_3 and HH_3] using all the above wavelet filters. Feature sets are created using the approximation and detail coefficients of each decomposed level.

V. SUPPORT VECTOR MACHINE

Support vector machine (SVM) have proven to have good performance in handwritten character recognition problems and is considered to be the state-of-the-art tool for linear and non-linear classification [19]. It belongs to the class of supervised learning algorithms, based on statistical learning theory. The SVM classifier has been originally proposed for binary classification in literature and learning algorithm comes from an optimal separating hyper-plane, developed by Vapnik [20].

For binary classification, a linear decision function $f(x)=w^Tx+b$ is used, where w is the weight vector and b is a bias. Classification is given by sign of f(x), which can be -1 or +1. The optimal solution is obtained when this hyper plane is located in the middle of the two classes (Fig. 4) and the points that constrain the width of the margin are called support vectors.

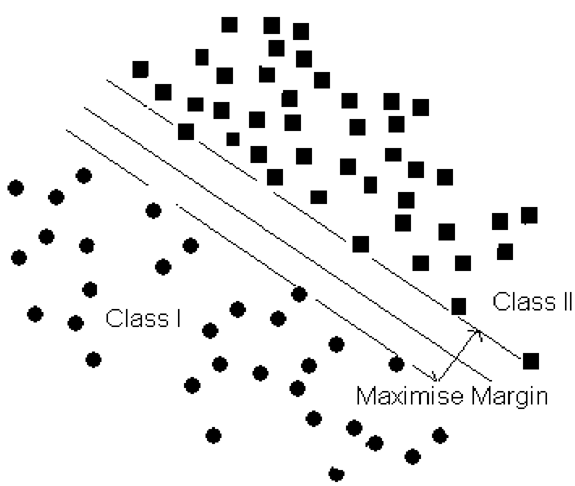

Figure. 4 Maximum margin hyper plane

Given a training set of instance – label pair $(X_i, y_i), i = 1,...,l$, where $X_i \in R^n$ and $y_i \in \{1,-1\}^l$, the support vector require the solution of the following optimization problem [19].

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i$$

subject to $(y_i w^T \phi(X_i) + b) \geq 1 - \xi_i$, $\xi_i \geq 0,$ (5)

where w is the weight vector, C is the soft margin parameter, ξ_i is a slack variable and b is the bias term. In the case of linearly inseparable feature space, the training vectors X_i are mapped into a higher dimensional space by the function ϕ. $K(X_i, X_j) \equiv \phi(X_i)^T \phi(X_j)$ is termed as the kernel function.

We use radial basis function (RBF) kernel

$$K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2), \gamma > 0 \qquad (6)$$

The binary SVM can be extended to multiclass [21]. Multiclass SVMs are usually implemented by combining several two-class SVMs either by one-versus-all method or one-versus-one method. In our problem, as the feature space is linearly inseparable, it is mapped into a high dimensional space through radial basis function kernel, so that the problem becomes linearly separable.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

For classification, support vector classifier with radial basis function kernel is selected. The parameter γ and C are set to 0.02 and 100 respectively. Stratified 10 fold cross-validation is used for all the experiment to report the classification accuracy, in which the dataset is divided into 10 folds and testing is repeated on each fold while using the remaining 9 folds for training. Evaluation of the proposed method is carried out by using the commonly known data mining metrics of classification rate that are described as follows[22,23]:

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP})$$
$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN})$$
$$\text{F - measure} = 2(\text{Precision.Recall})/(\text{Precision} + \text{Recall})$$
$$\text{TP rate} = \text{TP}/(\text{TP} + \text{FP})$$
$$\text{FP rate} = TN/(TN + FP)$$
$$ClassificationAccuracy = (TP + TN)/(TP + TN + FP + FN)$$
$$ErrorRate = (FP + FN)/(TP + TN + FP + FN) \qquad (7)$$

True positive rate is the proportion of actual positives which are predicted positive and true negative rate is the proportion of actual negative which are predicted negative. Precision or positive predictive value is defined as the proportion of predicted positives which are actual positive. Recall is defined as the proportion of the actual positives which are predicted positive and F-measure is the harmonic mean between precision and recall.

In order to test the efficiency of the proposed method, we have used 747 instances of 25 handwritten Malayalam characters segmented from a camera captured document. Each character in the database is size normalized to 64 x 64 pixels. Experimental results reveal that the smoothing and contrast enhancement produces an improvement in accuracy of about 3%. So these pre-processing steps are applied to each character image. Subsequently, wavelet decomposition is applied to yield

© 2010, IJARCS All Rights Reserved

four 32x32 sub images at LL_1, LH_1, HL_1 and HH_1 in the first level decomposition. The LL_1 sub-band is further decomposed into four 16 x 16 sub images, [LL_2, LH_2, HL_2 and HH_2], the LL_2 sub-band is decomposed into four 8x8 sub images [LL_3, LH_3, HL_3 and HH_3]. Families of wavelet filters, namely, Haar, Daubechies (db2, db3, db4, db5, db6, db7, db8, db9, db10), symlet (sym2, sym3, sym4, sym5, sym6, sym7, sym8), coiflet (coif1, coif2, coif3, coif4, coif5) and biorthogonal (bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5,bior3.7, bior3.9, bior4.4, bior5.5, bior6.8) are used for conducting experiments. The approximation coefficients (LL_2) in the second level decomposition of all the wavelet filters are chosen as feature set I. Approximation coefficients (LL_3) combined with detail coefficients (HL_3,LH_3,HH_3) of the third level decomposition of all wavelet filters are created as feature set II. The approximation and detail coefficients in the fourth level (LL_3) combined with detail coefficients of third level of all wavelet filters are chosen as feature set III.

From the calculated wavelet features for each feature set, the classification was carried out using SVM classifier. There are basically two challenges, one is to choose the best feature set and the other is to choose the best wavelet family. Fig. 5-8, displays the classification accuracies of different wavelet families. By analysing these figures, one can understand that, for feature set I, bior1.1 (93.12%), bior1.3 (91.23%), bior2.2 (89.34%), bior3.1 (91.90%), db2 (92.04%), coif1 (90.15%) and sym2 (92.04%) are giving good result and for feature set II, bior1.1(91.90%), bior1.3 (90.96%), bior1.5 (90.68%), bior2.2 (90.42%), bior2.4 (89.47%), bior4.4 (90.01%), bior5.5 (89.07%), db2 (90.96%), db3 (90.68%), sym2 (90.95%), sym3 (90.69%), sym4 (90.01%), sym5 (90.28%), sym7 (89.6%), sym8 (89.07%), coif1 (90.42%) and coif2 (90.96%) are all yield fine performance. For feature set III, bior1.1 (90.55%), coif1 (87.99%), db2 (88.80%), db3 (87.72%), sym2 (88.99%) and sym4 (88.39%) are better. While comparing the performances of different wavelet families, bior1.1 performed well for all the sets of features. Feature set I and feature set II outperformed feature set III for majority of experiments. From fig. 5-8, the best performing wavelet is picked up and the result is depicted in Table 1. The highest recognition accuracy of 93.12% with bior1.1 wavelet is obtained with feature set I. Even though the maximum accuracy is obtained with feature set I, performances of most of the wavelets are good with feature set II.

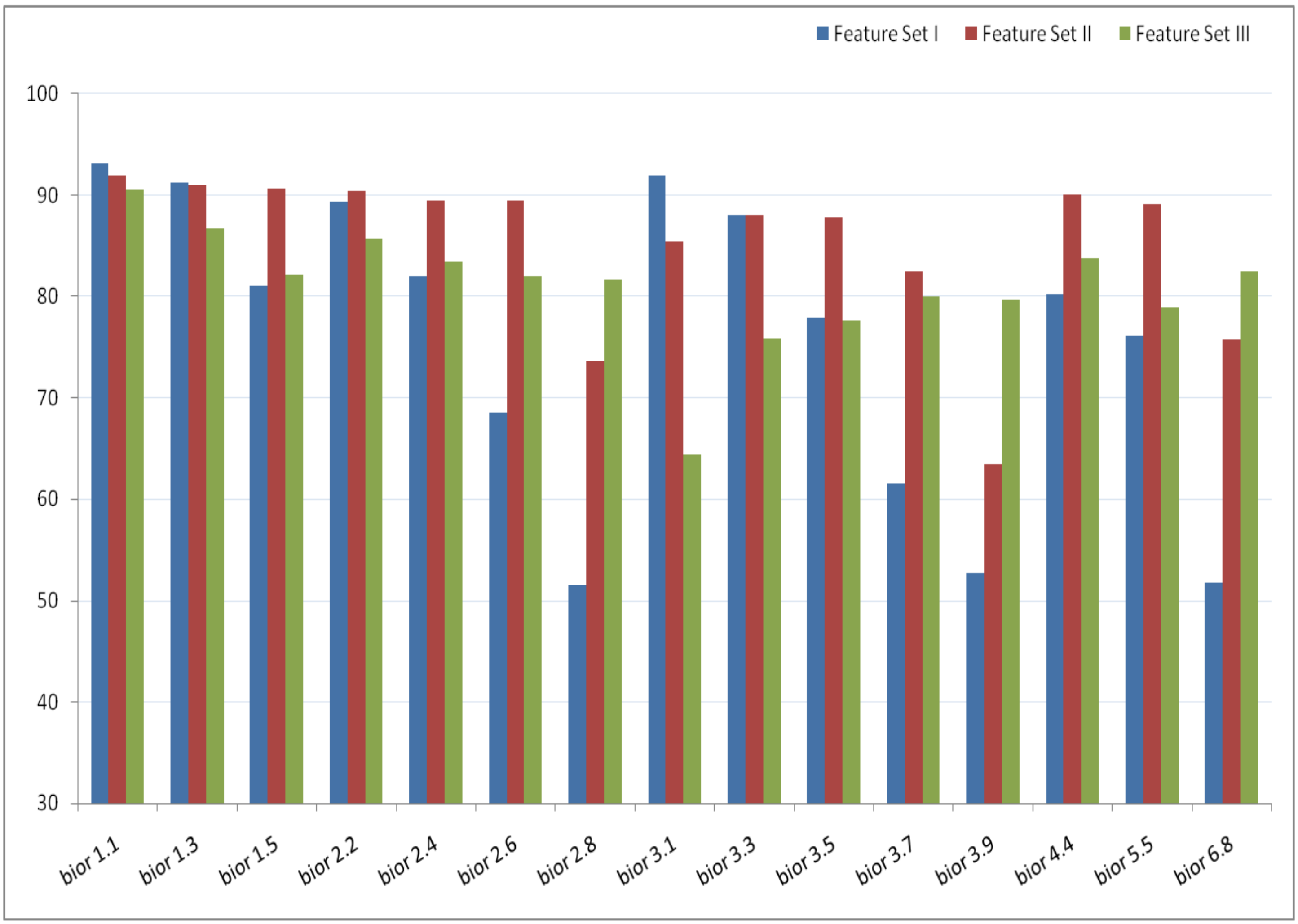

Figure. 5 Classification performance of Bi-orthogonal wavelet family for all feature sets

© 2010, IJARCS All Rights Reserved

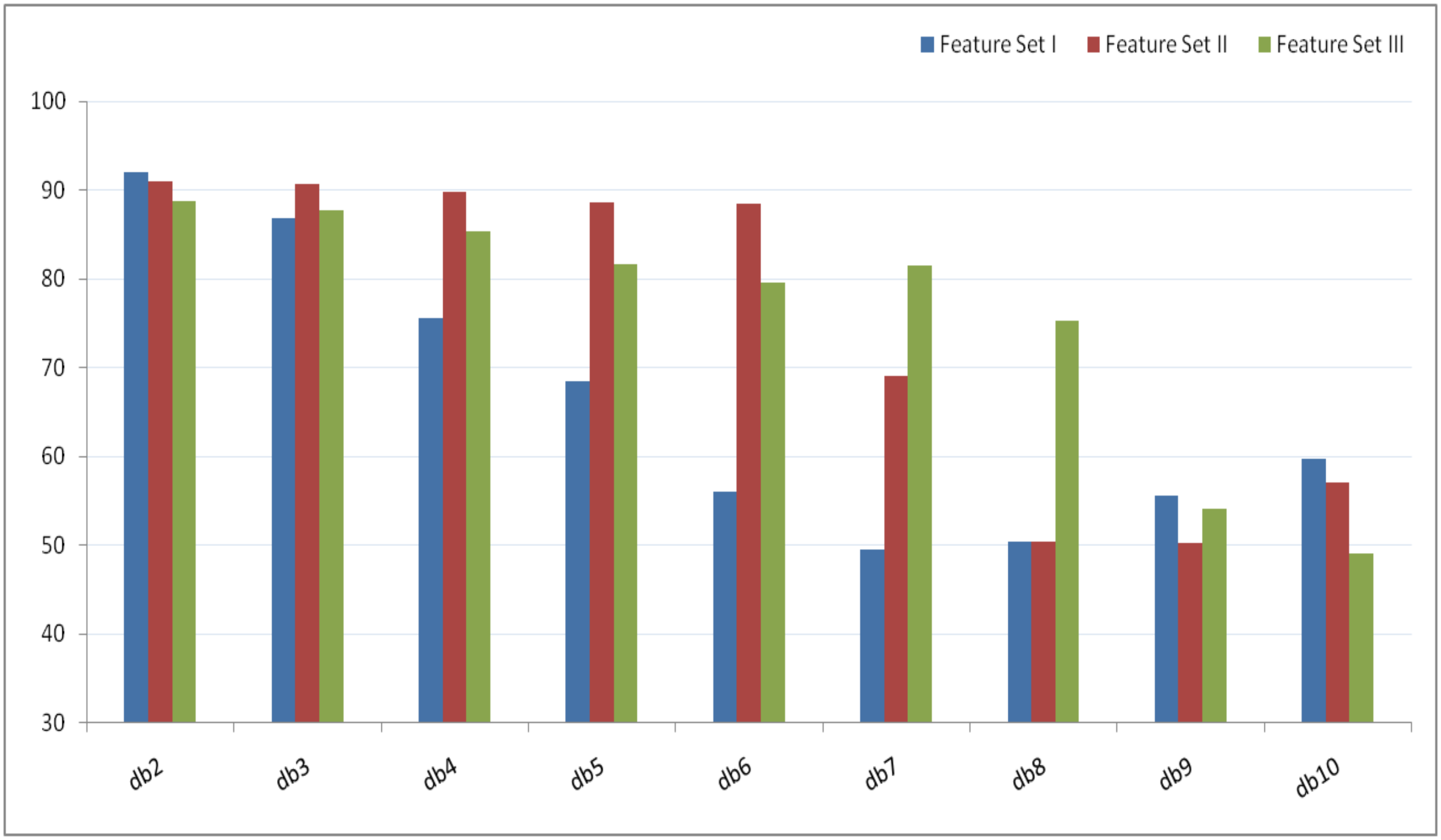

Figure. 6 Classification performance of Daubechies wavelet family for all feature sets

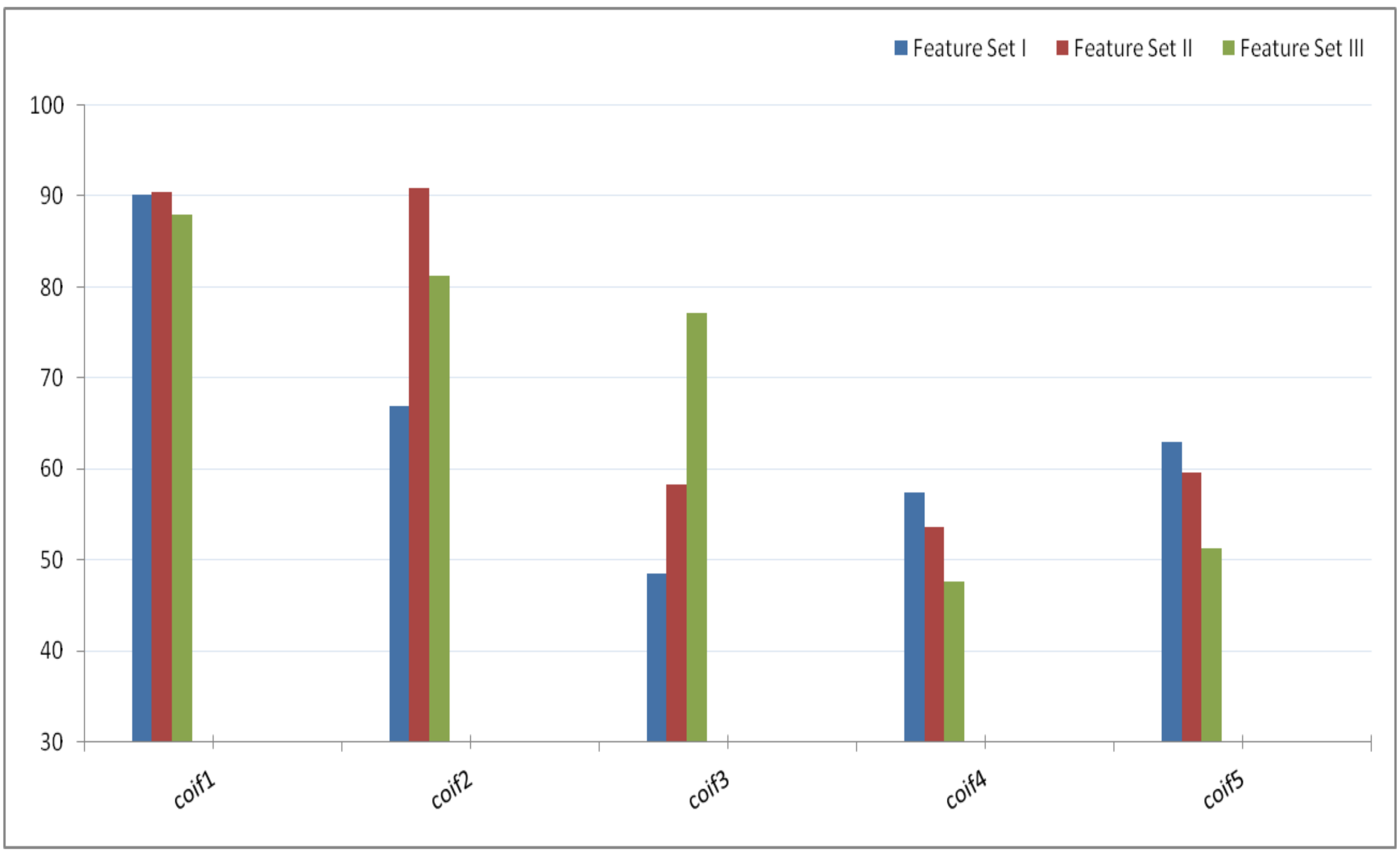

Figure. 7 Classification performance of Coiflet wavelet family for all feature sets

© 2010, IJARCS All Rights Reserved

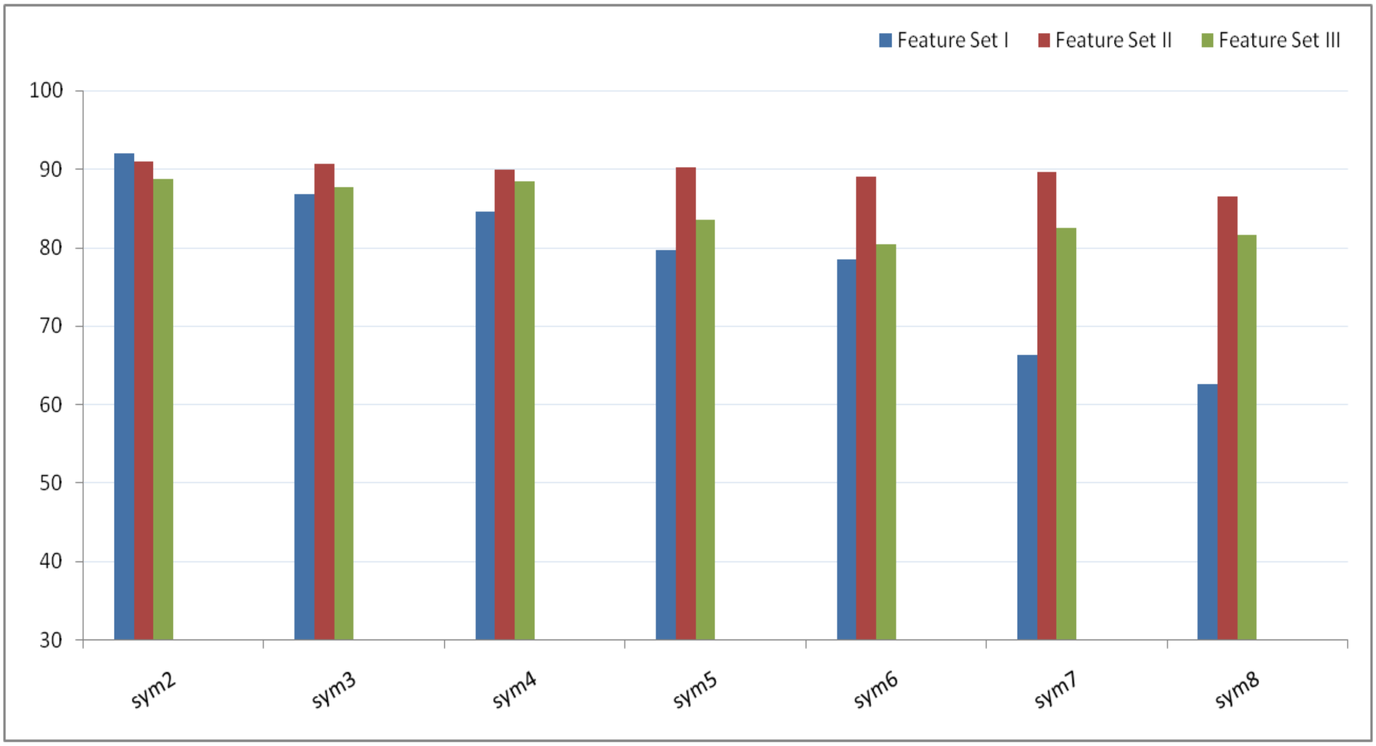

Figure. 8 Classification performance of Symlet wavelet family for all feature sets

Table 1: Recognition result of best performing filters

Wavelet filter	Feature set	Classification Accuracy	TP Rate	FP Rate	Precision	Recall	F-measure
bior 1.1	I	93.1174	0.931	0.003	0.933	0.931	0.931
bior 1.3	I	91.2281	0.912	0.003	0.914	0.912	0.913
bior 1.5	II	90.6883	0.907	0.004	0.91	0.907	0.907
bior 3.1	I	91.9028	0.919	0.003	0.92	0.919	0.919
db2	I	92.0378	0.92	0.003	0.922	0.92	0.921
db3	II	90.6883	0.907	0.004	0.911	0.907	0.908
coif 1	II	90.4184	0.904	0.004	0.908	0.904	0.905
coif 2	II	90.9582	0.91	0.004	0.911	0.91	0.909
sym 2	I	92.0378	0.92	0.003	0.922	0.92	0.921
sym3	II	90.6883	0.907	0.004	0.911	0.907	0.908

VII. CONCLUSION

In this paper, we propose a character recognition methodology for handwritten character images captured through mobile phone camera. The efficiency of the method depends on the pre-processing phase, segmentation of individual characters, feature extraction technique based on discrete wavelet transform and support vector machine classifier. Classification results reveal that the feature extraction method and classifier is discriminant. Our future research is focused on reducing the complexity of the system

© 2010, IJARCS All Rights Reserved

so as to integrate the system completely within a mobile platform.

VIII. REFERENCES

[1] Plamondon, R., & Srihari, S. N. (2000). Online and off-line handwriting recognition: a comprehensive survey. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(1), 63-84.

[2] Doermann, D., Jian, L., & Huiping, L. Progress in camera-based document image analysis. In Document Analysis and Recognition, 2003. Proceedings. Seventh International Conference on, 3-6 Aug. 2003 2003 (pp. 606-616 vol.601)

[3] Ghosh, D., Dube, T., & Shivaprasad, A. P. (2010). Script Recognition;A Review. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(12), 2142-2161, doi:10.1109/tpami.2010.30.

[4] Yuh-Rau Wang, W.-H. L., Shi-Jinn Horng (2011). A sliding window technique for efficient licence plate localization. Expert Systems with Applications, 38, 3142-3146.

[5] Wunsch, P., & Laine, A. F. (1995). Wavelet descriptors for multiresolution recognition of handprinted characters. Pattern Recognition, 28(8), 1237-1249, doi:10.1016/0031-3203(95)00001-g.

[6] Lee, S.-W., Kim, C.-H., Ma, H., & Tang, Y. Y. (1996). Multiresolution recognition of unconstrained handwritten numerals with wavelet transform and multilayer cluster neural network. Pattern Recognition, 29(12), 1953-1961, doi:10.1016/s0031-3203(96)00053-2.

[7] Chen, G. Y., Bui, T. D., & Krzyzak, A. (2003). Contour-based handwritten numeral recognition using multiwavelets and neural networks. Pattern Recognition, 36(7), 1597-1604, doi:10.1016/s0031-3203(02)00252-2.

[8] Liu C L, N. K., Sako H, Fujisawa H Handwritten digit recognition: Benchmarking of state-of-the-art techniques. Pattern Recognition, 36(10), 2271-2285

[9] Bellili A, G. M., Gallinari P (2003). An MLP-SVM combination architecture for offline handwritten digit recogniton:reduction of recogniton errors by support vector machine rejection mechanisms. International Journal of Document Analysis and Recogniton, 5, 244-252.

[10] Shanthi N, D. K. (2010). A novel SVM-based handwritten Tamil character recognition system. Pattern Analysis and Applications, 13, 173-180, doi:10.1007/s10044-009-0147-0.

[11] Bhowmik T K, G. P., A. Roy, Parui S. K (2009). SVM-based hierarchical architectures for handwritten Bangla character recognition. International Journal of Document Analysis and Recognition.

[12] John, J., Pramod, K. V., & Balakrishnan, K. (2012). Unconstrained Handwritten Malayalam Character Recognition using Wavelet Transform and Support vector Machine Classifier. Procedia Engineering, 30(0), 598-605, doi:10.1016/j.proeng.2012.01.904.

[13] Laine, M., & Nevalainen, O. S. A Standalone OCR System for Mobile Cameraphones. In Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on, 11-14 Sept. 2006 2006 (pp. 1-5)

[14] Koga, M., Mine, R., Kameyama, T., Takahashi, T., Yamazaki, M., & Yamaguchi, T. Camera-based Kanji OCR for mobile-phones: practical issues. In Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on, 29 Aug.-1 Sept. 2005 2005 (pp. 635-639 Vol. 632)

[15] Bae, K. S., Kim, K. K., Chung, Y. G., & Yu, W. P. Character recognition system for cellular phone with camera. In Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International, 26-28 July 2005 2005 (Vol. 1, pp. 539-544 Vol. 532)

[16] Kuo, J.-H., Huang, C.-M., Liao, W.-H., & Huang, C.-C. (2011). HuayuNavi: A Mobile Chinese Learning Application Based on Intelligent Character Recognition

[17] Daubechies, I. (1994). Ten lectures on wavelets (Vol. 61): SIAM.

[18] Mallat, S., G (1989). A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 674-693.

[19] Christopher, B., J, C, (1998). A tutorial on support vector machines for pattern recognition. Data Mining and Knowlege discovery, 2(2), 121-167.

[20] N, V. Vapnik. (1999). The Nature of Statistical Learning Theory. Berlin: Springer, Information science and statistics.

[21] Hsu, C., W,, & Lin, C., J, (2002). A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks, 2(13), 415-425.

[22] Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification taks. Inf. Process. Manage., 45(4), 427-437.

[23] Demsar, J. (2006). Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research, 7, 1-30.

© 2010, IJARCS All Rights Reserved