Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Engineering 30 (2012) 598 – 605
Procedia Engineering
www.elsevier.com/locate/procedia

International Conference on Communication Technology and System Design 2011

Unconstrained Handwritten Malayalam Character Recognition using Wavelet Transform and Support vector Machine Classifier

Jomy John[a], Pramod K. V.[b], Kannan Balakrishnan[c] a*

Department of Computer Applications, Cochin University of Science and Technology, Kochi, 682 022, India

Abstract

This paper presents the application of wavelet processing in the domain of handwritten character recognition. To attain high recognition rate, robust feature extractors and powerful classifiers that are invariant to degree of variability of human writing are needed. The proposed scheme consists of two stages: a feature extraction stage, which is based on Haar wavelet transform and a classification stage that uses support vector machine classifier. Experimental results show that the proposed method is effective.

Keywords: handwritten character recognition; Haar wavelet transfor; support vector machine; RBF kernel;

1. Introduction

Wavelet transforms are efficient tool for character recognition. It decomposes an image of a character 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.

The area of character recognition has been receiving considerable attention due its versatile range of application domain including postal automation, bank check processing, automating of processing of large volumes of data, language based learning, ledgering catalogue for library, reading aid for blind etc. Even though sufficient study has been proposed for languages like Chinese, Latin and English, research on Indian script is still active and demanding.

Malayalam is one of twenty two scheduled languages in India, with rich literary heritage. Character recognition in Malayalam language is more complex due to enormously large character set and high

* Jomy John. Tel.: +0-944-643-8725; fax: +0-000-000-0000 .
E-mail address: jomyeldos@gmail.com

1877-7058 © 2011 Published by Elsevier Ltd.
doi:10.1016/j.proeng.2012.01.904

similarity between characters. The problem becomes more difficult in the handwritten domain due to the varying writing style of each individual. The Malayalam script consists of 15 vowels and 36 consonants. This paper investigates the use of wavelet transform for the recognition of handwritten Malayalam characters.

Powerful classifiers that are invariant to degree of variability of human writing are needed for a large class problem. A standard approach of character recognition is to train the classifier to predict one of the output classes. The support vector machines (SVMs) are effective discriminative classifiers with good generalization and convergence property and it had proven to have good performance in handwritten character recognition problems.

Wunsch and Laine [1] used wavelet features extracted from contour of the handwritten characters for classification using neural networks. Lee et al. [2] used wavelet features extracted from hand written numerals are classified it using multilayer cluster neural network. Chen et al. [3] developed multi-wavelet descriptor for contour of handwritten numerals using neural network. But all these works deal with only few classes as opposite to the present large class classification problem. In this study, all the forty four basic Malayalam characters have been considered. In a similar study, [4] used zero crossings of wave packets to classify twenty classes using feed forward back propagation network. To our best knowledge, the use of SVM in offline handwritten Malayalam characters represents a novelty.

The paper is organized as follows: the next section introduces wavelet theory and the proposed method for feature extraction. Section 3 discusses SVM, Section 4 evaluates performance of the system and Section 5 concludes the paper.

2. Image features extracted from Haar wavelets

2.1. Wavelet Theory

The space frequency localization and multi-resolution analysis capability of a wavelet makes it an efficient tool in analysing images. In this paper, Haar wavelets [5] have been used for multi-resolution feature extraction. This wavelet was introduced by Hungarian mathematician Alfred Haar in 1910 and it is one of the earliest wavelet with low computing requirements, which is also known as a compact orthonormal wavelet transform.

The Haar scaling function $\varphi(x)$ and the Haar wavelet function $\psi(x)$ are as follows:

$$\varphi(x) = \begin{cases} 1 & 0 \le x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad \psi(x) = \begin{cases} 1 & 0 \le x < ½ \\ -1 & ½ \le x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Discrete wavelet transform can be obtained using the analysis filters for decomposition and the synthesis filters for reconstruction. As we are interested in obtaining the features for classification purpose, we are dealing with only the analysis filter. The scaling function $\varphi(x)$ and the wavelet function $\psi(x)$ associated with the scaling filter h_φ and the wavelet filter h_ψ are:

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

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

In two-dimensional wavelet decomposition, the analysis scaling function can be written as the product of two one-dimensional scaling functions $\varphi(x)$ and $\varphi(y)$.

$$\varphi(x,y) = \varphi(x)\varphi(y)$$

If $\psi(x)$ is the one-dimensional wavelet associated with the scaling function, then, the three two-dimensional analysis wavelets are defined as:

$$\psi^H(x,y) = \psi(x)\varphi(y)$$

$$\psi^V(x,y) = \varphi(x)\psi(y)$$

$$\psi^D(x,y) = \psi(x)\psi(y)$$

where $\psi^H(x,y)$, $\psi^V(x,y)$ and $\psi^D(x,y)$ correspond to horizontal, vertical and diagonal wavelets respectively.

The multi-resolution technique can be implemented using sub-band decomposition in which the image of a character is decomposed into wavelet coefficients [6]. 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. 1 explains the decomposition, in which j+1 stands for the starting scale, m and n are row and column directions.

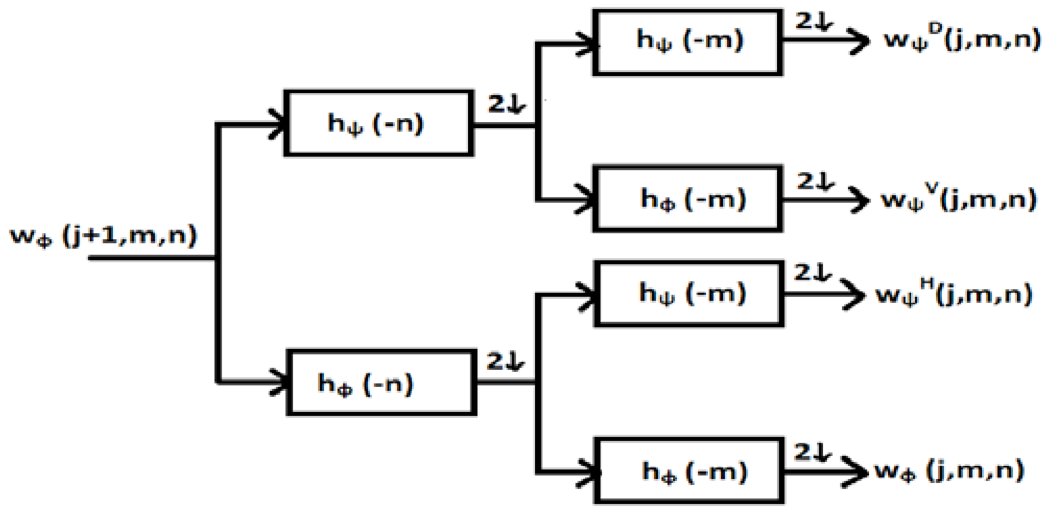

Fig. 1 Decomposition using analysis filter bank

2.2. Malayalam Character Segmentation

As the collected samples contain text lines of characters, segmentation is required to isolate each character to form the database. Segmentation is based on projection analysis and connected component labeling. The steps for character segmentation are as follows:

2.2.1. Line Segmentation

Text lines 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.

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

From this profile, the peak-valley points are identified and that is used for line separation.

2.2.2. Character Segmentation

Each extracted line is segmented into a sequence of isolated characters using connected component labeling algorithm. The minimum bounded rectangle containing the component are extracted and stored in the database.

2.3. 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. The original image is first converted to gray scale and then size normalized to 64x64 pixels. The Haar wavelet decomposition with Haar analysis filter h_φ = [0.707107, 0.707107] and h_ψ= [0.707107, -0.707107] are applied to each character 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}. The character image 'ah' and its third level decomposition are displayed in Fig.2.

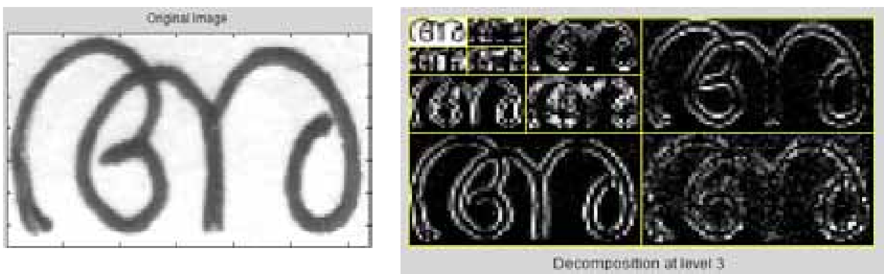

Fig. 2 Original image and decomposition at level 3 of character 'ah'

3. Support vector machine classification

SVM is one of the popular techniques for pattern recognition and is considered to be the state-of-the-art tool for linear and non-linear classification [7]. 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 [8].

Given a training set of instance – label pair $(x_i, y_i), i = 1,\ldots,l$, where $x_i \in R^n$ and $y_i \in \{1,-1\}^l$, the support vector require the solution of the following optimization problem [7].

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

$$\textit{subject to } (y_i w^T \phi(X_i) + b) \geq 1 - \xi_i$$

$$\xi_i \geq 0,\ i = 1,\ldots,l$$

where 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. The kernel function is termed as:

$$K(X_i, X_j) \equiv \phi(X_i)^T \phi(X_j).$$

Commonly used kernel functions are:

Linear: $K(X_i, X_j) = x_i^T x_j$ *Polynomial*: $K(X_i, X_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$

Radial basis function: $K(X_i, X_j) = \exp(-\gamma \|x_i - x_j\|^2, \gamma > 0$

Sigmoid: $K(X_i, X_j) = \tanh(\gamma x_i^T x_j + r)$*, where* γ, r, d *are kernel parameters.*

The effectiveness of SVM depends on the selection of kernel, the kernel's parameters, and soft margin parameter C. The binary SVM can be extended to multiclass [9]. 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.

4. Performance evaluation

In the experiment, we have used MATLAB 7.8.0 for wavelet decomposition and feature extraction of images, and WLSVM [10] which is an implementation of LibSVM [11] running under Weka environment for classification.

Data were collected from different persons of the population in Kerala, including different age groups without imposing any constraints. It represents wide variety of writing styles. Digitization of collected samples are done by a Flat-bed scanner (manufactured by HP, Model Name: Scanjet 2400), by setting dpi to 300. The experiment had been carried out on a database of 10,000 handwritten isolated Malayalam characters written by 228 different writers. It contains all the 44 basic characters. Fig. 3 displays samples of all these characters.

Fig. 3 Samples of each character from the database

In order to demonstrate the effectiveness of the proposed method, two types of experiments have been considered. In the first experiment, only approximation of Haar wavelet coefficients at decomposition level 3 (LL_3 subband) are chosen and in the second experiment, decomposition on level 2 (LL_2 subband) have considered. As support vector machine with radial basis function (RBF) kernel has better performance in character recognition problems, it has been chosen for the present study. We set the parameters $\gamma=[2^4, 2^3, 2^2, \ldots, 2^{-10}]$ and $C=[2^{12}, 2^{11}, 2^{10}, \ldots, 2^{-2}]$ for coarse estimation and on fine tuning, we obtained the optimum result with $\gamma = 0.02$, C = 100. For all experiments, the database was split with a random process into training (80%) and testing (20%).

The steps involved can be summarized as follows:

1. Convert the character image into gray scale.
2. Scale the image so that it fits exactly into a 64 x 64 matrix.
3. Perform the 2-D Haar wavelet transform on the scaled image.
4. Train the SVM with the extracted feature vectors from the training dataset.
5. Test the SVM to obtain the recognition rates

Classification result is 89.64% using third level decomposition. Experiments on second level decomposition show that 90.25% of classes are correctly classified. Individual results for all the classes on the second level decomposition are displayed in Table 1.

Table 1. Classification result

Class Id	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
1	0.947	0	1	0.947	0.973	0.974
2	1	0.003	0.909	1	0.952	0.999
3	0.831	0.006	0.831	0.831	0.831	0.913
4	0.796	0.007	0.736	0.796	0.765	0.894
5	0.923	0.002	0.9	0.923	0.911	0.961
6	0.897	0.004	0.833	0.897	0.864	0.947
7	0.943	0.002	0.917	0.943	0.93	0.971
8	0.875	0.001	0.955	0.875	0.913	0.937
9	0.894	0.001	0.955	0.894	0.923	0.946
10	0.833	0.003	0.854	0.833	0.843	0.915
11	0.944	0.002	0.927	0.944	0.936	0.971
12	0.875	0.002	0.933	0.875	0.903	0.937
13	0.962	0.002	0.943	0.962	0.952	0.98
14	0.936	0.002	0.936	0.936	0.936	0.967
15	0.929	0.002	0.945	0.929	0.937	0.964
16	0.943	0.003	0.846	0.943	0.892	0.97
17	0.909	0.002	0.882	0.909	0.896	0.954
18	0.878	0.002	0.915	0.878	0.896	0.938
19	0.951	0.001	0.967	0.951	0.959	0.975
20	0.974	0.002	0.927	0.974	0.95	0.986
21	0.946	0.003	0.875	0.946	0.909	0.972
22	1	0.002	0.9	1	0.947	0.999
23	0.896	0.002	0.935	0.896	0.915	0.947
24	0.907	0.002	0.907	0.907	0.907	0.952
25	0.868	0.001	0.958	0.868	0.911	0.933
26	0.933	0.004	0.857	0.933	0.894	0.965
27	0.878	0.002	0.9	0.878	0.889	0.938
28	0.925	0.003	0.881	0.925	0.902	0.961
29	0.945	0.003	0.897	0.945	0.92	0.971
30	0.872	0.003	0.872	0.872	0.872	0.935
31	0.846	0.003	0.846	0.846	0.846	0.922
32	0.93	0.004	0.851	0.93	0.889	0.963
33	0.907	0.004	0.848	0.907	0.876	0.952
34	0.868	0.004	0.805	0.868	0.835	0.932
35	0.889	0.001	0.952	0.889	0.92	0.944

36	0.977	0.001	0.977	0.977	0.977	0.988
37	0.911	0.003	0.891	0.911	0.901	0.954
38	0.953	0.002	0.932	0.953	0.943	0.976
39	0.805	0.001	0.971	0.805	0.88	0.902
40	0.804	0.002	0.911	0.804	0.854	0.901
41	0.822	0.003	0.86	0.822	0.841	0.91
42	0.783	0.003	0.857	0.783	0.818	0.89
43	0.886	0.002	0.929	0.886	0.907	0.942
44	0.953	0.001	0.976	0.953	0.965	0.976
Weighted Average	0.903	0.002	0.904	0.903	0.902	0.95

5. Conclusion

The wavelet coefficients of an image have multi-resolution representation of original image. The coarse resolution wavelet coefficients normally represent the overall shape of the image, while the fine resolution coefficients represent the details of the image. We use the Haar wavelet features at different resolution scales in the experiments reported here. Support vector machine with RBF kernel is used for classification. In this 44-class classification problem, the recognizer yields a result of 90.25% classification accuracy.

References

[1] P Wunsch, A F Laine. “Wavelet descriptors for multiresolution recognition of handprinted characters”. Pattern Recognition, 1995, 28(8): 1237–1249

[2] S W Lee, C H Kim, H Ma, Y Y Tang. “Multiresolution recognition of unconstrained handwritten numerals with wavelet transform and multilayer cluster neural network”. Pattern Recognition, 1996, 29: 1953–1961

[3] G.Y.Chen, T.D.Bui, A.Krzyzak. “Contour based handwritten numeral recognition using multiwavelets and neural networks”, Pattern Recognition, 36(7)(2003)1597-1604

[4] G. Raju and K. Revathy.”Wavepackets in the Recognition of Isolated Handwritten Characters”, Proceedings of the World Congress on Engineering 2007 Vol I WCE 2007, July 2 - 4, 2007, London, U.K.

[5] R. S. Stankovic, B. J. Falkowski. ”The Haar wavelet transform: its status and achievements”, Computers and Electrical Engineering 29 (2003) 25–44

[6] S.G. Mallat. “A theory for multiresolution signal decomposition:The wavelet representation”. IEEE Trans. on PatternAnalysis and Machine Intelligence, 11(7):674–693, 1989.

[7] J C Christopher Burges (1998) “A tutorial on support vector machines for pattern recognition”. Data Min. Knowl. Discov. 2(2): 121-167

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

[9] C W Hsu, CJ Lin (2002) “A comparison of methods for multiclass Support Vector Machines”. IEEE Trans. Neural Netw. 13(2), 415–425

[10] Yasser EL-Manzalawy and Vasant Honavar, “WLSVM : Integrating LibSVM into Weka Environment”, 2005

[11] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : “A library for support vector machines”. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011.