Global and Local Features for Recognition of Online Handwritten Numerals and Tamil Characters

A. G. Ramakrishnan*
Indian Institute of Science
Bangalore, India
ramkiag@ee.iisc.ernet.in

Bhargava Urala K
Indian Institute of Science
Bangalore, India
bhargava@mile.ee.iisc.ernet.in

ABSTRACT

Feature extraction is a key step in the recognition of online handwritten data and is well investigated in literature. In the case of Tamil online handwritten characters, global features such as those derived from discrete Fourier transform (DFT), discrete cosine transform (DCT), wavelet transform have been used to capture overall information about the data. On the hand, local features such as (x, y) coordinates, n^{th} derivative, curvature and angular features have also been used. In this paper, we investigate the efficacy of using global features alone (DFT, DCT), local features alone (preprocessed (x, y) coordinates) and a combination of both global and local features. Our classifier, a support vector machine (SVM) with radial basis function (RBF) kernel, is trained and tested on the IWFHR 2006 Tamil handwritten character recognition competition dataset. We have obtained more than 95% accuracy on the test dataset which is greater than the best score reported in the literature. Further, we have used a combination of global and local features on a publicly available database of Indo-Arabic numerals and obtained an accuracy of more than 98%.

Categories and Subject Descriptors

I.7.5 [**Document Capture**]: Optical Character Recognition—*online handwriting recognition*

Keywords

Online Handwriting Recognition, Tamil Character Recognition, Numeral Recognition, Feature Extraction, Pattern Recognition, Discrete Fourier Transform, Discrete Cosine Transform, Support Vector Machine

1. INTRODUCTION

One of the earliest works in literature that deals with representation of Tamil characters is by Sundaresan and Keerthi [16]. Here, the authors report that the performance of angle features with a neural network classifier was poor and that wavelet features were superior to Fourier features in capturing the overall information as well increasing interclass separability when used with neural networks. Bharath and Madhvanath in [6] use normalised (x, y) features, first derivatives, directional features and angular features for recognition of Tamil words using hidden Markov models (HMM). In their work in [7], Dinesh and Sridhar introduce a novel feature for recognition called *star* feature and compare it with normalised (x, y) coordinates and directional features (described in [11]) on Tamil and Roman characters and on Indo-Arabic numerals.

In other Indic languages, usage of subspace features along with principal component analysis (PCA) for recognition of Devanagari characters is reported in [10]. Rajkumar et al. in [13] have experimented with Fourier, wavelet and Hilbert features among others along with an SVM classifier and have achieved high recognition rates for Telugu characters. Arora and Namboodiri propose a hybrid model comprising many features such as preprocessed (x, y) coordinates, magnitude of Fourier coefficients, n^{th} order moments, length, area and curvature. The results of using these features along with different classifiers on Telugu and Malayalam characters are reported in [5]. It must be noted that the recognition primitives in this case are individual strokes that are later concatenated to form words or isolated characters using a post-processing methodology. Whereas, in our work, the recognition primitives are the so called *stroke groups* or *symbols* which are described in [14] and [15].

This paper is organised as follows. Section 2 briefly describes the training and testing dataset used in our work. Section 3 describes in detail the features we have worked with - local, global and a combination of the two. We conclude in Section 4 with the tabulation of results, some discussion and suggestions for future work. Samples of the handwritten Tamil symbols used for training our classifier are displayed at the end of our paper.

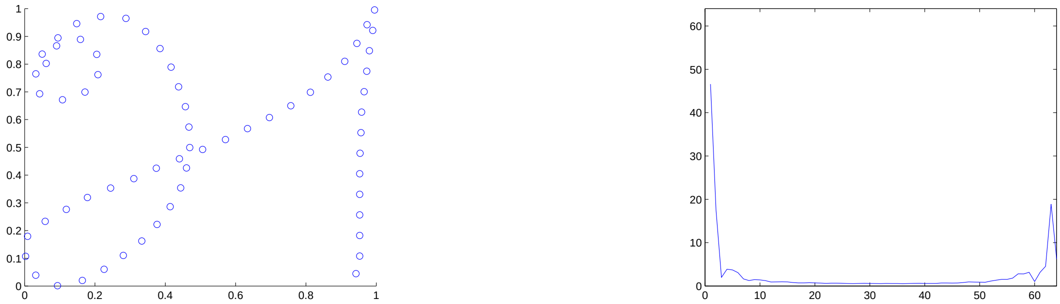
2. DATASET DESCRIPTION

We have used the HP Labs Tamil online handwritten character dataset which has been made publicly available for research [9] and was created for a competition in IWFHR 2006. It consists of 50,385 training samples and 26,926 test samples across 156 symbols or stroke groups. However the 156^{th} class in the dataset is a combination of already present classes 10 and 28 and is therefore neglected. We use the rest of the 155 classes as recognition primitives in our classifier.

*Corresponding author

MOCR '13, August 24, 2013, Washington DC, USA

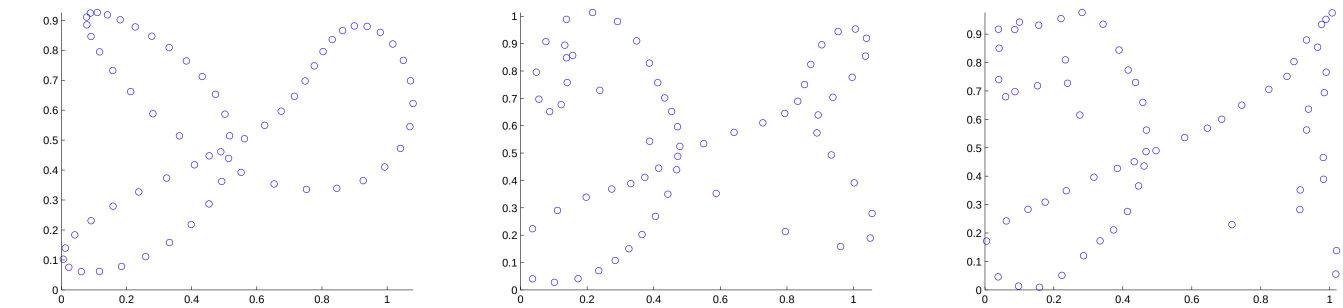
http://dx.doi.org/10.1145/2505377.2505390.



(a) Preprocessed character (vowel /a/)

(b) Magnitude of its DFT

Figure 1: A sample preprocessed character and the magnitude of its DFT



(a) From 8 point complex-valued Fourier coefficients

(b) From 16 point complex-valued Fourier coefficients

(c) From 32 point complex-valued Fourier coefficients

Figure 2: Reconstruction of the character from the truncated Fourier transform coefficients

We have used the Ethem Alpaydin handwritten Indo-Arabic numeral database which consists of ten classes (0 - 9), about 750 training samples and about 350 test samples per class. This database is publicly available, consists of multiple representations of the numerals and in [4], the authors have investigated the use of multiple classifiers for recognition of these handwritten digits.

3. GLOBAL AND LOCAL FEATURES

As mentioned in the abstract we have experimented with three different types of features.

- Local features only: Using the HP training dataset, we trained a classifier on the preprocessed (x, y) features and performed recognition on the test dataset.

- Global features only: We used discrete Fourier transform (DFT) and discrete cosine transform (DCT) of the preprocessed feature vector of each symbol to train the classifier. We have also investigated the effect of truncating the Fourier transform feature vector.

- Combination of local and global features: We trained the classifier on feature vectors obtained by concatenating the local features and truncated Fourier features. We also experimented with using first derivative features along with the previously mentioned features.

3.1 Local Features

The data corresponding to every symbol is preprocessed before feature extraction in the online handwriting recognition workflow. The steps followed for preprocessing are - normalisation, smoothing and resampling. Preprocessing is done to reduce the effect of noise, remove redundant points in the data, account for size variations and reduce dependency on writing styles [6, 15, 5]. The preprocessing procedure is the same as described in [15] except that the number of points of the preprocessed vector is taken to be 64, which results in a feature vector length of 128. SVM classifier with RBF kernel is trained on the preprocessed data from the IWFHR database and a grid search with 5-fold cross validation is performed to find ideal training parameters. Cross validation and test accuracies are shown in Table 1. We have used LIBSVM - an open source implementation of SVM [2].

3.2 Global Features

The discrete Fourier transform is used to extract global features from the handwritten data. We treat the preprocessed (x, y) coordinates as a 64-point complex-valued vector and then take its Fourier transform, i.e

$$F(k) = \sum_{i=0}^{N-1} z_i e^{-j2\pi ik/N} \tag{1}$$

where, $N = 64$ and $z_i = x_i + jy_i$. This approach is inspired from use of Fourier descriptors which are used in image processing for object recognition and retrieval applications [8]. Note that a 64 point complex-valued vector would need a 128 length feature vector to represent it.

The Fourier transform lends itself to compression because typically most of the energy of the vector is contained in very few coefficients. This can be observed in Figure 1(b). We experimented with truncating the feature vector F to 8, 16 and 32 complex points and then attempting to reconstruct the original symbol by taking the inverse Fourier transform of the truncated feature vector. The results of reconstruction for a symbol are shown in Figures 2 (a), (b) and (c) for truncation to 8, 16, and 32 complex coefficients respectively. It can be seen that truncating to 32 complex points and subsequent inverse transform produces an acceptable reconstruction.

The discrete cosine transform (DCT) was also considered as a possible global feature. DCT is obtained using,

$$C(k) = \sum_{i=0}^{N-1} w_k z_i \cos \frac{\pi k(2i+1)}{2N} \tag{2}$$

where, $N = 64$, $z_i = x_i + jy_i$, $w_k = \sqrt{1/N}$, for $k = 0$ and $w_k = \sqrt{2/N}$, for $1 \leq k \leq N-1$.

A separate SVM classifier was trained with the vectors F and C computed from the training samples of each symbol in the IWFHR dataset. An RBF kernel was used and a grid search with 5-fold cross validation was performed to find optimal training parameters. This procedure was repeated after truncating the feature vector F to a 32 point complex valued vector. Cross validation and test accuracies for all the cases are shown in Table 1.

3.3 Combined Features

Global features are good at capturing overall information about the data. However, they generally do not work well with similar classes that have minor point-wise variations. On the other hand, local features are extracted at each point and therefore result in a good inter-class separation. To obtain the best of both worlds, we experimented with the concatenation of global and local features. Preprocessed (x, y) coordinates (from Section 3.1) and truncated Fourier descriptors (from Section 3.2) were concatenated and an SVM with RBF kernel was trained. First derivative features were computed and concatenated to the above vector and was used to train an SVM with RBF kernel. The first derivative features were computed using the equations [12],

$$d_x(i) = \frac{(x_i - x_{i-1}) + \frac{(x_{i+1} - x_{i-1})}{2}}{2}, 2 \leq i \leq 63 \tag{3}$$

and

$$d_y(i) = \frac{(y_i - y_{i-1}) + \frac{(y_{i+1} - y_{i-1})}{2}}{2}, 2 \leq i \leq 63 \tag{4}$$

As we can see from the above equations, the value of first and last points of d_x and d_y cannot be computed. Therefore, we make $d_x(1) = d_x(2)$, $d_y(1) = d_y(2)$, $d_x(64) = d_x(63)$ and $d_y(64) = d_y(63)$. A grid search with 5-fold cross validation was used to obtain the best training parameters for both cases and the accuracies of cross-validation and testing on the test dataset are tabulated in Table 1.

Table 1: Cross Validation and Test Accuracies for different features on the complete test set of IWFHR 2006 online Tamil handwritten character database

L: Length of feature vector, CVA: Cross validation accuracy, TA: Test accuracy

Feature	L	CVA (%)	TA (%)
(x,y)	128	91.8	92.46
DFT	128	91.7	95.78
DCT	128	91.6	93.84
Truncated DFT	64	91.5	95.69
(x,y)+ truncated DFT	192	91.3	95.85
(x,y)+ truncated DFT+first derivative	320	91.5	95.86

Table 2: Comparison of Tamil character recognition with best results in the literature

TA: Test accuracy

	Feature	TA (%)
IWFHR 1^{st} place [1]	Online+ Offline	93.53
IWFHR 2^{st} place [1]	Online	91.2
IWFHR 3^{st} place [1]	Online	90.72
Dinesh and Sridhar	Star [7] features	80.2
AGR and Bhargava	(x,y)+ truncated DFT+first derivative	95.86

Table 3: Comparison of numeral recognition with best results in the literature

TA: Test accuracy

	Database	TA (%)
Verma et al.	UNIPEN	98.2
AGR and Bhargava	Alpaydin database	98.14
Alimoglu and Alpaydin	Alpaydin database	95.26
Uchida et al.	Alpaydin	89.4
Ahmad and Maen		~ 95

4. RESULTS AND DISCUSSION

Table 1 lists feature vector length, cross validation accuracies during classifier training as well as testing accuracies for each of the different feature extraction methods mentioned in the previous sections. The test dataset consisted of 26926 test samples across the 156 original symbols. However, after removing test samples of one symbol (156^{th} class) from this dataset, for reasons mentioned in Section 2, the total number of test samples reduces to 26754.

We can see from Table 1 that using Fourier descriptor features offers a significant improvement over using local features alone (92.46% to 95.78%) whereas improvement due to usage of DCT features is lesser (92.46% to 93.84%). The best entry for Tamil handwritten character recognition competition organised as part of IWFHR 2006 reports an accuracy of 93.53% [1], having used a combination of online and offline features. The best accuracy reported for using online features only is reported to be 91.2%. The test accuracies for all the features we have experimented, except preprocessed (x, y) coordinates alone, are greater than the best result of the competition. Table 2 lists some of the other test accuracies reported in literature.

Our choice of truncating the Fourier coefficient vector to 32 complex points instead of 64 is justified when we see that the test accuracies of the two methods do not differ by a significant amount. Though we expect that the increase in test accuracy of the combined features compared to truncated DFT features be higher, we find that it is only 0.16%. It is interesting to note that the 5-fold cross validation accuracy is approximately the same ($\sim$ 91%) for all the choices of feature extraction methods.

4.1 Results on online numeral database

The best result in literature on online handwritten numeral data is by Verma et al. in [18] who have obtained 98.2% on the UNIPEN numeral database using neural network classifier. Our work combining global (truncated DFT) and local ((x, y) coordinates) features to train an SVM classifier resulted in an accuracy of 98.14% on the Ethem Alpaydin dataset. On the same database, Uchida et al. using dynamic time warping obtain 89.4% accuracy, whereas Alimoglu and Alpaydin report an accuracy of 95.26% [17, 4]. Ahmad and Maen, in [3] use a finite transition network to obtain an accuracy of 95% on their own database of about 3000 samples. A summary of these results can be seen in Table 3.

In future, we would like to apply these feature extraction techniques to tackle the problem of online handwritten word recognition and see which method of extraction would work best. It would also be interesting to study the confusion matrix obtained by using each of these methods for character recognition. Since we have obtained high recognition accuracies (>95%) from some methods, it could also be worthwhile adopting this approach to other languages as well.

5. ACKNOWLEDGMENTS

We thank all research students and project staff at MILE lab, IISc. We would like to specially thank Technology Development for Indian Languages (TDIL), Department of Information Technology, Government of India for funding this work.

6. REFERENCES

[1] *IWFHR 2006 - Online Tamil handwritten character recognition competition.* http://algoval.essex.ac.uk:8080/iwfhr2006/index.jsp.

[2] *LIBSVM - A Library for Support Vector Machines.* http://www.csie.ntu.edu.tw/cjlin/libsvm/.

[3] A. Ahmad and H. Maen. Recognition of online handwritten Arabic digits using structural features and transition network. In *Informatica*, pages 275–281, 2008.

[4] F. Alimoglu and E. Alpaydin. Combining multiple representations and classifiers for pen-based handwritten digit recognition. In *Proc. ICDAR*, volume 2, pages 637–640, 1997.

[5] A. Arora and A. M. Namboodiri. A hybrid model for recognition of online handwriting in Indian scripts. In *Proc. ICFHR*, pages 433–438, 2010.

[6] A. Bharath and S. Madhvanath. Hidden Markov Models for online handwritten Tamil word recognition. In *Proc. IWFHR*, pages 506–510, 2007.

[7] M. Dinesh and M. K. Sridhar. A feature based on encoding of relative position of a point in the character for online handwritten character recognition. In *Proc. ICDAR*, pages 1014–1017, 2007.

[8] R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* Addison Wesley and Longman Publishing, 1992.

[9] HP Labs India. *Isolated Tamil Handwritten Character Dataset.* http://www.hpl.hp.com/india/research/penhw-interfaces-1linguistics.html.

[10] N. Joshi, G. Sita, A. G. Ramakrishnan, V. Deepu, and S. Madhvanath. Machine recognition of online handwritten Devanagari characters. In *Proc. ICDAR*, volume 2, pages 1156–1160, 2005.

[11] N. Kato and S. Omachi. A handwritten character recognition system using directional element feature and asymmetric Mahalanobis distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(3):258–262, March 1999.

[12] E. Keogh and M. Pazzani. Derivative dynamic time warping. In *First SIAM International Conference on Data Mining*, 2001.

[13] J. Rajkumar, K. Mariraja, K. Kanakapriya, S. Nishanthini, and V. S. Chakravarthy. Two schemas for online character recognition of Telugu script based on Support Vector Machines. In *Proc. ICFHR*, pages 565–570, 2012.

[14] S. Sundaram and A. G. Ramakrishnan. Attention-feedback based robust segmentation of online handwritten isolated Tamil words. *ACM Transactions on Asian Language Information Processing*, In press, March 2013.

[15] S. Sundaram, B. Urala, and A. G. Ramakrishnan. Language models for online handwritten Tamil word recognition. In *Proc. Workshop on Document Analysis and Recognition*, 2012.

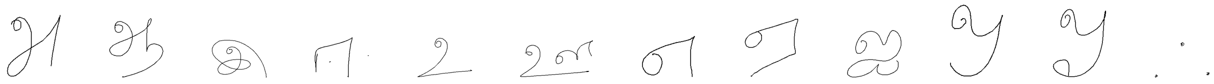
[16] C. S. Sundaresan and S. S. Keerthi. A study of representations for pen based handwriting recognition of Tamil characters. In *Proc. ICDAR*, pages 422–425, 1999.

[17] S. Uchida, M. Fukotomi, K. Ogawara, and Y. Feng. Non-Markovian dynamic time warping. In *Proc. ICPR*, pages 2294–2297, 2012.

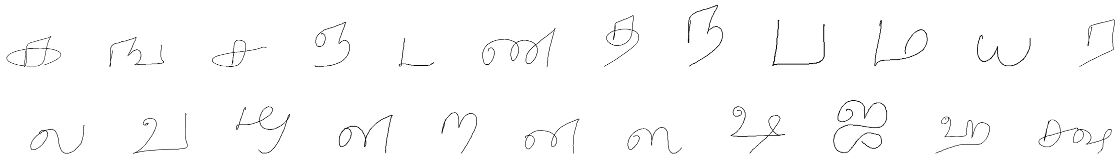
[18] B. Verma, J. Lu, M. Ghosh, and R. Ghosh. A feature extraction technique for online handwriting recognition. In *Proc. IJCNN*, volume 2, pages 1337–1341, 2004.

Samples from the Tamil training dataset

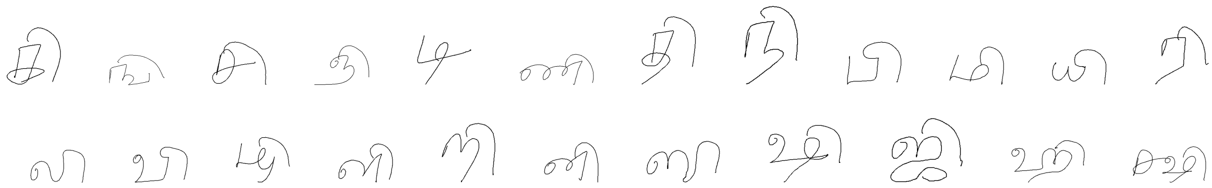
Vowels



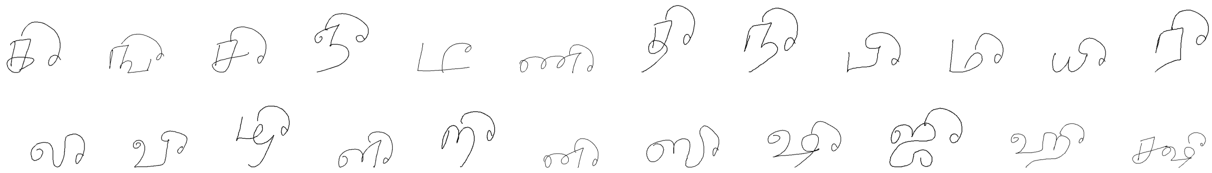
Consonants



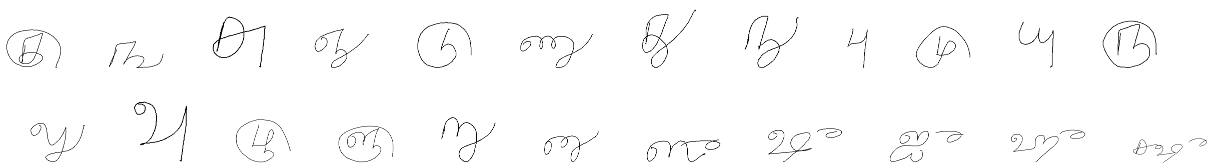
CV combinations with /i/



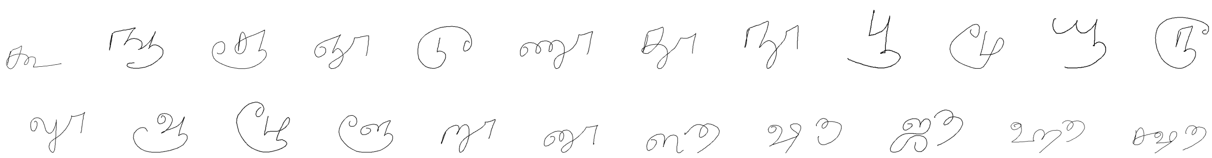
CV combinations with /I/



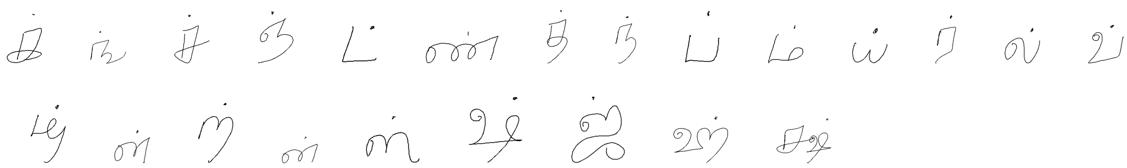
CV combinations with /u/



CV combinations with /U/



Pure consonants



Vowel modifiers for /A/, /e/, /E/, /ai/ and /shri/