Comparative Study of Devnagari Handwritten Character Recognition using Different Feature and Classifiers

U. Pal[1], T. Wakabayashi[2], F. Kimura[2]
[1]Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata-108, India
[2]Graduate School of Engineering, Mie University, TSU, Mie 514-8507, Japan
Email:umapada@isical.ac.in

Abstract

In recent years research towards Indian handwritten character recognition is getting increasing attention. Many approaches have been proposed by the researchers towards handwritten Indian character recognition and many recognition systems for isolated handwritten numerals/characters are available in the literature. To get idea of the recognition results of different classifiers and to provide new benchmark for future research, in this paper a comparative study of Devnagari handwritten character recognition using twelve different classifiers and four sets of feature is presented. Projection distance, subspace method, linear discriminant function, support vector machines, modified quadratic discriminant function, mirror image learning, Euclidean distance, nearest neighbour, k-Nearest neighbour, modified projection distance, compound projection distance, and compound modified quadratic discriminant function are used as different classifiers. Feature sets used in the classifiers are computed based on curvature and gradient information obtained from binary as well as gray-scale images.

1. Introduction

Recognition of handwritten characters has been a popular research area for many years because of its various application potentials [1]. Some of its potential application areas are Postal Automation, Bank cheque processing, automatic data entry, etc. In recent years research towards Indian handwriting character recognition is getting increasing attention. Many approaches have been proposed by the researchers towards handwritten Indian character recognition and many recognition systems for isolated handwritten numerals/characters are available in the literature [2,3].

India is a multi-lingual multi-script country and there are twenty two languages. Eleven scripts are used to write these languages and Devnagari is the most popular script in India. First research report on handwritten Devnagari characters was published in 1977 [4] but not much research work is done after that. At present researchers have started working on handwritten Devnagari characters. Many research reports are available towards Devnagari numeral recognition [5,6] but to the best of our knowledge there are only a few research reports available on Devnagari off-line handwritten character recognition [7-9] after the year 1977.

To get idea of the recognition results of different classifiers and to provide new benchmark for future research, in this paper a comparative study of Devnagari handwritten character recognition results is reported here. To compare the performance, twelve different classifiers and four different features computed from gradient and curvature information of the binary as well as gray-scale images are used here. Classifier like Projection distance (PD), Subspace method (SM), Linear discriminant function (LDF), Support vector machines (SVM), Modified quadratic discriminant function (MQDF), Mirror image learning (MIL), Euclidean distance (ED), Nearest neighbour, k-Nearest neighbour (k-NN), Modified Projection distance (MPD), Compound projection distance (CPD), and Compound modified quadratic discriminant function (CMQDF) are considered.

Rest of the paper is organized as follows. In Section 2 properties of Devnagari script are discussed. Feature extraction procedures are reported in Section 3. In Section 4, we briefly explain different classifiers used for performance analysis. Details comparative results are discussed in Section 5. Finally, conclusion is given in Section 6.

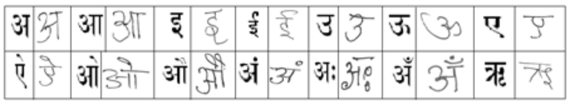
2. Properties of Devnagari script

Devnagari is the most popular script in India and the Indian national language, Hindi, is written in Devnagari script. Nepali, Sanskrit and Marathi are also written in Devnagari script. Moreover, Hindi is the third most popular language in the world [2].The alphabet of the modern Devnagari script consists of 14 vowels and 33 consonants. These characters are called *basic characters*. The shape of basic characters of Devnagari script are shown in Fig.1 and we used these 47 basic characters for our experiment. Writing mode in Devnagari script is from left to right. The concept of upper/lower case is absent in Devnagari script. In Devnagari script a vowel following a consonant takes a modified shape. Depending on the vowel, its modified

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.244

IEEE computer society

shape is placed at the left, right (or both) or bottom of the consonant. These modified shapes are called *modified characters*. A consonant or vowel following a consonant sometimes takes a compound orthographic shape, which we call as *compound character*. Compound characters can be combinations of two consonants as well as a consonant and a vowel. Compounding of three or four characters also exists in the script. There are about 280 compound characters in Devnagari [2].



(a)

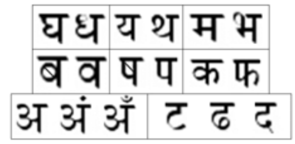
(b)

Figure 1: Samples of handwritten Devnagari basic characters (a) Vowels (b) Consonants. To get an idea about the shape of the character, samples of printed characters are shown in the left side of the respective handwritten characters.

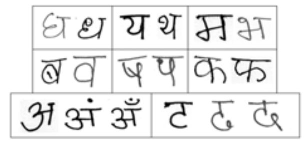
The complexity of a handwritten character recognition system increases mainly because of various writing styles of different individuals. Most of the errors in such system arise because of the confusion among the similar shaped characters. In Devnagari there are many similar shaped characters. Examples of some groups of similar shaped characters are shown in Fig.2. To get an idea of similar shaped printed as well as handwritten characters, we provide the samples of both printed and handwritten Devnagari characters in Fig.2. Although there are some differences between the samples of a group in the printed characters but the difference in the corresponding handwritten samples is very less. From the Fig.2(b) it can be seen that shapes of two or more characters of a group is very similar due to handwritten style of different individuals and such shape similarity is the main reason of recognition errors.

3. Feature extraction

Four sets of features (two sets from binary and two sets from gray-scale images) are used for recognition purpose. One set of features is based on gradient and other set is computed from both gradient and curvature information. Our data set was grey-scale and to get the feature on binary images we converted the grey-scale image into binary using Otsu method [10]. Dimension of all feature set is 392 and the computation methods of the feature sets are given as follows. In our past experiment [9] we obtained improved results in 392 dimension hence we have considered 392 dimensional feature for this comparative study.



(a)



(b)

Figure 2: Examples of some similar shaped Devnagari characters. (a) Printed samples (b) corresponding handwritten samples of (a).

3.1 Computation of gradient feature

To get gradient feature, at first, a 2 x 2 mean filtering is applied 4 times on the input image and a non-linear size normalization is done on the image [11]. Here the image is normalized into 148 x 148 pixels and this size is decided from the experiment. Normalized image is then segmented into 49 x 49 blocks. Compromising trade-off between accuracy and complexity, this block size is decided. To obtain 49 x 49 blocks from a pattern of 148 x 148 pixels we used k= (49/148)*(i-1)+1 and l=(49/148)*(j-1)+1, where (i, j) is the coordinates of 148x148 patterns and (k, l) is the coordinates of 49x49 blocks.

A Roberts filter is then applied on the normalized image to obtain gradient image. Next, the arc tangent of the gradient (direction of gradient) is initially quantized into 32 directions with $\pi/16$ intervals and the strength of the gradient is accumulated with each of the quantized direction. By strength of gradient (SG) we mean $SG = \sqrt{(\Delta u)^2 + (\Delta v)^2}$, and by direction of gradient ($\theta(x, y)$) we mean $\theta(x, y) = \tan^{-1}\frac{\Delta v}{\Delta u}$, where $\Delta u = f(x+1, y+1) - f(x, y)$, and $\Delta v = f(x+1, y) - f(x, y+1)$. Here $f(x, y)$ is a gray scale at (*x, y*) point. Finally, using Gaussian like filter 49 x 49 blocks are down samples into 7 x 7 blocks and 32 directional frequencies are down sampled into 8 directions and as a results we get 392 (7x7x8) dimensional feature vector. For details of gradient feature see the paper [12].

3.2 Computation of curvature feature

Curvature features can be computed in different ways and we computed curvature feature using bi-quadratic interpolation method because it gave better results according to the experiment of Shi et al. [12]. To get the features following steps are executed.

Step 1: The direction of gradient is quantized to 32 levels as discussed above.
Step 2: The curvature c is computed by bi-quadratic interpolation method and quantized into 3 levels using a threshold t (for concave, linear and convex regions). For concave region $c \leq -t$, for linear region $-t < c < t$ and for convex region $c \geq t$. We assume t as 0.15 in our experiment.
Step 3: The strength of the gradient is accumulated in each of the 32 directions and in each of the 3 curvatures levels of each 49x49 blocks to get 49x49 local joint spectra of directions and curvatures.
Step 4: A spatial and directional resolution is made as follows. A smoothing filter [1 4 6 4 1] is used to get 16 directions from 32 directions. On this resultant image, another smoothing filter [1 2 1] is used to get 8 directions from 16 directions. Further more, we use a 31 x 31 two-dimensional Gaussian-like filter to get smoothed 7 × 7 blocks from 49 x 49 blocks. So, we get 7×7×8 = 392 dimensional feature vector. Using curvature feature in 3 levels we get 392 × 3 =1176 dimensional features.
Step 5: Using principal component analysis we reduce 1176 dimensional feature vector to 392 dimensional feature vector and we fed this 392 dimensional feature vector to our classifiers for comparison.

4. Brief description of the classifiers

Twelve different classifiers like projection distance, subspace method, linear discriminant function, support vector machines, modified quadratic discriminant function, mirror image learning, Euclidean distance, nearest neighbour, k-Nearest neighbour, modified projection distance, compound projection distance and compound modified quadratic discriminant function are used for comparative study. Both parametric and non-parametric classifiers are used for our experiment. Detail descriptions of these classifiers can be obtained in the literature [18] and hence we are not giving their details here. However, some of the classifiers are briefly discussed as follows.

Euclidian Distance (ED): The Euclidean distance between the input pattern and the mean vector is defined by

$$g_l^2(X) = \|X - M_l\|^2$$

where X is the input feature vector of size (dimensionality) n, M_l is the mean vector of class l. The input vector is classified to such class l^* that minimizes the Euclidean distance. Hereafter the subscript l denoting the class is omitted for the sake of simplicity.

Projection Distance (PD): The projection distance is defined by

$$g_{pd}^2(X) = \|X - M\|^2 - \sum_{i=1}^{k} \{\Phi_i^T (X - M)\}^2$$

and gives the distance from the input pattern X to the minimum mean square error hyperplane that approximates the distribution of the sample, where Φ_i denotes the i-th eigenvector of the covariance matrix, and k is the dimensionality of the hyperplane as well as the number of the dominant eigen vectors ($k < n$). When $k = 0$ the projection distance reduces to the Euclidean distance.

Subspace method (SM): For a bipolar distribution on a spherical surface with $\|X\| = 1$ the mean vector M is a zero vector ($M = 0$) because the distribution is symmetric in respect to the origin. Then the projection distance for the distribution is given by

$$g^2(X) = 1 - \sum_{i=1}^{k} \{\Phi_i^T X\}^2$$

where Φ_i is the i-th eigenvector of the autocorrelation matrix. The second term of the above expression is used as the similarity measure of CLAFIC (Class Featuring Information Compression) and the subspace method [14].

Modified Quadratic Discriminant Function (MQDF): Modified quadratic discriminant function is defined as follows [13].

$$g(X) = (N + N_0 + n - 1)\ln[1 + \frac{1}{N_0\sigma^2}[\|X - M\|^2 - \sum_{i=1}^{k} \frac{\lambda_i}{\lambda_i + \frac{N_0}{N}\sigma^2}\{\Phi_i^T(X - M)\}^2]] + \sum_{i=1}^{k} \ln(\lambda_i + \frac{N_0}{N}\sigma^2)$$

where X is the feature vector of an input character; M is a mean vector of samples; Φ_i^T is the i^{th} eigen vector of the sample covariance matrix; λ_i is the i^{th} eigen value of the sample covariance matrix; k is the number of eigen values considered here, n is the feature size; σ^2 is the initial estimation of a variance; N is the number of learning samples; and N_o is a confidence constant for σ.

Modified Projection Distance (MPD): The modified projection distance is defined by

$$g^2(X) = \|X - M\|^2 - \sum_{i=1}^{k} \frac{(1-\alpha)\lambda_i}{(1-\alpha)\lambda_i + \alpha\sigma^2} \{\Phi_i^T (X - M)\}^2$$

where α is a parameter which takes [0, 1]. When $\alpha = 0$, this classifier gives the same value as that of Projection Distance. When $\alpha = 1$, this gives the same value as that of Euclidian Distance. The value of α we used here is decided by preliminary experiment.

Linear Discriminant Function (LD): Linear discriminant function is defined by

$$g\ (X) = W^T X + W_0$$
$$W\ = S_w^{-1} M$$
$$W_0\ = -\frac{1}{2} M^T S_w^{-1} M$$

where S_w is within-class covariance matrix.

Mirror Image Learning (MIL): Mirror Image Learning (MIL) is a corrective learning algorithm proposed to improve the learning effectiveness of class conditional distributions. The MIL generates a mirror image of a pattern which belongs to one of a pair of confusing classes to increases the size of the learning sample of the other class. For details of MIL see [14].

Compound Projection Distance (CPD) and Compound Modified Quadratic Discriminant Function (CMQDF):. Several compound discriminant functions have been derived from the projection distance, the MQDF and so on to discriminate the similar shaped confusing character pairs effectively for Chinese character recognition [15]. The compound projection distance is an extended projection distance such that the difference between the mean vectors of the confusing pairs are directly taken into account, and is defined as a linear combination of the projection distance and the extension.

$$g_{cpd}^2(X) = (1-\delta) g_{pd}^2(X) + \delta G_{cpd}^2(X) \qquad (0 \le \delta \le 1)$$

The extension is defined by

$$G_{cpd}^2(X) = \frac{\left[M^T Y - \sum_{i=1}^{k} \{M^T \Phi_i\}\{Y^T \Phi_i\} \right]^2}{M^T M - \sum_{i=1}^{k} \{M^T \Phi_i\}^2}$$

$$M = M_2 - M_1,\ Y = X - M_1$$

Here M_1 and Φi are the mean vector and the i-th eigen vector of the covariance matrix of the one class respectively, and M_2 is the mean vector of the other class. Compound MQDF is the similar extension of the MQDF[15].

Support Vector Machine (SVM):
An SVM is defined for two-class problem and it finds the optimal hyper-plane which maximizes the distance, the *margin*, between the nearest examples of both classes, named *support vectors* (SVs). Given a training database of M data: $\{x_m | m=1,...,M\}$, the linear SVM classifier is then defined as:

$$f(x) = \sum_j \alpha_j x_j \cdot x + b$$

where $\{x_j\}$ are the set of support vectors and the parameters α_j and b have been determined by solving a quadratic problem [16].

The linear SVM can be extended to a non-linear classifier by replacing the inner product between the input vector x and the SVs x_j, to a kernel function K defined as: $K(x, y) = \phi(x) \cdot \phi(y)$. This kernel function should satisfy the Mercer's Condition [16]. There are many kernels and in our work we have used Gaussian kernel because it gave highest performance in our experiment. We used LIBSVM (www.csie.ntu.edu.tw/~cjlin/libsvm) for our experiment.

5. Result and discussions

Data used for the present work were collected from different individuals. We tested 36172 samples of Devnagari basic characters (vowels as well as consonants) for the experiment of the proposed work. (The data set available for research purpose on email request to umapada@isical.ac.in). We have used 5-fold cross validation scheme for recognition result computation. Here database is divided into 5 subsets and testing is done on each subset using rest of the subsets for learning. The recognition rates for all the test subsets are averaged to calculate recognition accuracy.

Using twelve classifiers we computed different results obtained from both grey and binary images of Devnagari handwritten characters and the results are shown in Table 1.

Table 1: Detail results of different classifiers.

Classifier	Grey image Gradient (392 dim.)	Grey image Curvature (392 dim.)	Binary image Gradient (392 dim.)	Binary image Curvature (392 dim.)	Average
PD	92.76	93.76	92.77	93.53	93.21
SM	92.61	93.62	92.61	93.35	93.04
LD	86.76	89.02	86.78	88.88	87.86
SVM	93.38	94.52	93.59	94.36	93.96
MQDF	94.24	94.78	94.14	94.50	94.42
MIL	94.74	95.19	94.74	95.09	94.94
ED	77.94	80.06	77.89	80.08	78.99
NN	87.16	86.72	87.69	87.22	87.19
k-NN	89.85	89.78	90.06	89.96	89.91
MPD	93.92	94.44	93.83	94.44	94.15
CPD	94.45	94.97	94.33	94.75	94.62
CMQDF	94.39	94.92	94.32	94.62	94.56

From the experiment we noted that MIL classifier provided best results among all the 12 classifiers considered here. The MIL classifier gave an accuracy of 94.74% and 95.19% for 392 dim. gradient and curvature features computed from grey images, respectively. Again with respect to 392 dim. gradient and curvature features computed from binary images it gave an accuracy of 94.74% and 95.09%. More interestingly, MIL provided best results among all classifiers in all the respective features. The Euclidean Distance (ED) showed the lowest results (77.89%) among the classifiers. It gave a result of 77.94% and 80.06% in the case of 392 dimensional gradient grey and 392 dimensional curvature grey images, respectively. Again in 392 dimensional gradient binary images and 392 dimensional curvature binary images it gave a result of 77.89% and 80.08%, respectively.

From the experiment we observed that curvature feature provided higher results than gradient features in all the classifiers except NN and k-NN. NN and k-NN classifiers show slightly lower results in curvature features than gradient features. Also from the experiment we noticed that except ED, NN and k-NN classifiers the features computed in gray-scale images show better results than that of binary images.

A graphical representation of these results is also shown in Fig.3. If we ranked the 12 different classifiers considered here based on their overall recognition results then first 5 classifiers (in decreasing order) are MIL, CPD, CMQDF, MQDF and SVM. Average recognition results for four different features are given in the last column of Table 1.

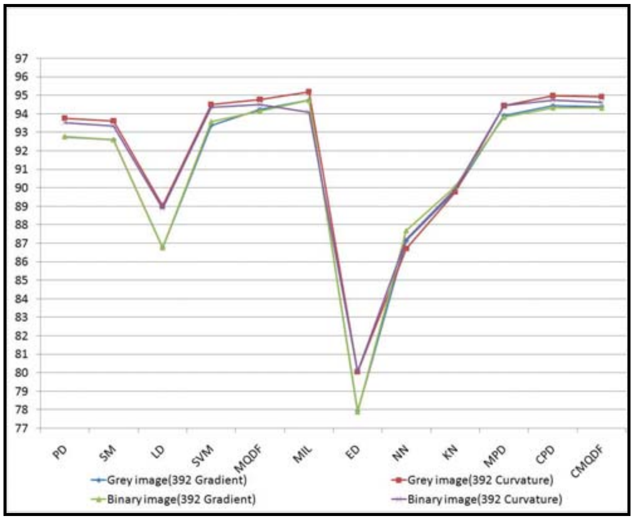


Figure 3: Graphical representation of results obtained from different classifiers.

To the best of our knowledge there exists only four pieces of work on off-line handwritten Devnagari characters and we compared our current best results with those of the existing pieces of work. Details comparative results are given in Table 2.

Table 2: Comparison of results.

Sl. no.	Method proposed by	Data size	Accuracy obtained
1.	Kumar and Singh [7]	200	80%
2.	Sharma et al.[8]	11270	80.36%
3.	Pal et al. [9]	36172	94.24%
4.	Pal et al. [17]	36172	95.13%
5.	Current method (by MIL classifier)	36172	95.19%

6. Conclusion

To get the idea of the recognition results of different classifiers and to provide new benchmark for future research, in this paper a comparative study of Devnagari handwritten character recognition using twelve different classifiers is reported here. Results of different classifiers are discussed and justifications of the results obtained are briefed. We noted that Mirror Image Learning gave overall better results among the classifiers and shown highest results (95.19%) accuracy on grey-scale curvature features. The authors hope this benchmark of results will be helpful to the researchers for future work.

7. References

[1] R. Plamondon and S. N. Srihari, “On-Line and off-line handwritten recognition: A comprehensive survey”, IEEE Trans on PAMI, Vol.22, pp.62-84, 2000.

[2] U. Pal and B. B. Chaudhuri, “Indian script character recognition: A Survey”, Pattern Recognition, Vol. 37, pp. 1887-1899, 2004.

[3] C. L. Liu and C. Y. Suen, “A new benchmark on the recognition of handwritten Bangla and Farshi numeral characters”, In Proc. 11th ICFHR, 2008.

[4] I. K. Sethi and B. Chatterjee, “Machine Recognition of constrained Hand printed Devnagari”, Pattern Recognition, Vol. 9, pp. 69-75, 1977.

[5] U. Pal, T. Wakabayashi, N. Sharma and F. Kimura, “Handwritten Numeral Recognition of Six Popular Indian Scripts”, In Proc. 9th ICDAR, pp. 749-753, 2007.

[6] M. Hanmandlu and O. V. R. Murthy, “Fuzzy Model Based Recognition of Handwritten Hindi Numerals”, In Proc. Intl. Conf. on Cognition and Recognition, pp. 490-496, 2005.

[7] S. Kumar and C. Singh, “A Study of Zernike Moments and its use in Devnagari Handwritten Character Recognition”, In Proc. Intl. Conf. on Cognition and Recognition, pp. 514-520, 2005.

[8] N. Sharma, U. Pal, F. Kimura and S. Pal, “Recognition of Offline Handwritten Devnagari Characters using Quadratic Classifier”, In Proc. Indian Conference on Computer Vision Graphics and Image Processing, pp- 805-816, 2006.

[9] U. Pal, N. Sharma, T. Wakabayashi and F. Kimura, “Off-Line Handwritten Character Recognition of Devnagari Script”, In Proc. 9th ICDAR, pp.496-500, 2007.

[10] N. Otsu, “A Threshold selection method from grey level histogram”, IEEE Trans on SMC, Vol.9, pp.62-66, 1979.

[11] F. Kimura, T. Wakabayashi, S. Tsuruoka and Y. Miyake, “Improvement of handwritten Japanese character recognition using weighted direction code histogram”, Pattern Recognition, Vol.30, No.8, pp.1329-1337, 1997.

[12] M. Shi, Y. Fujisawa, T. Wakabayashi, and F. Kimura, “Handwritten numeral recognition using gradient and curvature of gray scale images”, Pattern Recognition, Vol.35, pp.2051-2059, 2000.

[13] F. Kimura, K. Takashina, S. Tsuruoka and Y. Miyake, “Modified quadratic discriminant function and the application to Chinese character recognition”, IEEE Trans. on PAMI, Vol. 9, pp 149-153, 1987.

[14] T. Wakabayashi, M. Shi, W. Ohyama, and F. Kimura: "A Comparative Study on Mirror Image Learning and ALSM": In Proc. 8th IWFHR, pp. 151-156, 2002.

[15] T. Nakajima, T. Wakabayashi, F. Kimura, Y. Miyake, “Accuracy Improvement by Compound Discriminant Function for Resembling Character Recognition in Japanese], Trans. of IEICE, Vol.J83-D-II, pp.623-633, 2000.

[16] V. Vapnik, “The Nature of Statistical Learning Theory”, *Springer Verlang,* 1995.

[17] U. Pal, S. Chanda, T. Wakabayashi and F. Kimura, “Accuracy Improvement of Devnagari Character Recognition Combining SVM and MQDF”, In Proc. 11th ICFHR, pp.367-372, 2008.

[18] S. Meng, “Study on Accuracy Improvement of Handwritten Character Recognition” Ph.D. Thesis, Mie University, 2002.