International Journal of Computer Applications (0975 – 8887)
Volume 47– No.1, June 2012

Recognition of Handwritten Gurmukhi Numeral using Gabor Filters

Sukhpreet Singh
Renu Dhir
Department of Computer Science and Engineering
B.R. Ambedkar, PhD. NIT
Jalandhar, India

ABSTRACT
Isolated handwritten character recognition has been the subject of intensive research during the last decades because it is useful in wide range of real world problems. It also provides a solution for processing large volumes of data automatically. There is an emerging trend in the research of recognizing handwritten characters and numerals of many Indian languages and scripts. In this paper, two different feature sets based on Gabor filter have been used for recognition. One is being GABM having dimensionality 210 and other being GABN with dimensionality 200. The SVM classifier with RBF (Radial Basis Function) kernel is used for classification. The performance of the Gabor filter is tested on the database consisting of 1500 samples for basic 10 numerals of Gurmukhi script, collected from different writers. By using 7-fold cross validation, accuracy of 99.53% using second feature set and 98.4% using first feature set are observed. To obtain better results preprocessing of noise removal and normalization processes before feature extraction are recommended.

General Terms
OCR, Pattern Recognition.

Keywords
Gabor features, Gabor filters, Handwritten Gurmukhi numeral recognition, SVM with RBF kernel.

1. INTRODUCTION
OPTICAL Character Recognition (OCR) is one of the important tasks of machine learning. It is the process of converting the images of handwritten, typewritten or printed text (usually captured by a scanner) into machine-editable text or computer processable format, such as ASCII code. OCR has applications in postal code recognition, automatic data entry into large administrative systems, banking, automatic cartography, 3D object recognition, digital libraries, invoice and receipt processing, reading devices for blind and personal digital assistants. OCR includes essential problems of pattern recognition. Accuracy, flexibility and speed are the three main features that characterize a good OCR system. OCR aims at enabling computers to recognize optical symbols without human intervention. This is accomplished by searching a match between the features extracted from a given symbol's image and the library of image models. The basic process of OCR Systems is shown in Figure1.

During last two decades, many advances have been achieved in OCR area. In the field of document image analysis and recognition, researchers have achieved great success in character recognition.

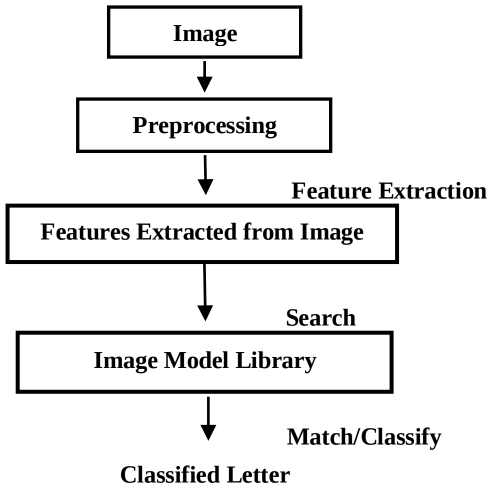


Figure 1: The Basic process of an OCR system

Recognition of handwritten characters has been a popular research area for many years because of its various application potentials, such as Postal automation, Bank cheque processing, Automatic data entry etc. However, there still exists challenging problem of handwritten character recognition [1]. Work is available towards postal automation of non-Indian language documents and sorting systems are also available for postal automation in several countries like USA, UK, Japan, Germany etc. But no such sorting system is available for Indian post. System development towards postal automation for a country like India is difficult because of its multi-lingual and multi-script behaviour.

Some pieces of work have been done on the recognition of Indian handwritten numerals [2–6] but not much work has been reported for Gurmukhi Language [7–9].

Dharamveer Sharma, et al. [7] have used structural and statistical features to recognize extracted Gurmukhi digits from Gurmukhi documents. They tested their technique on both Roman and Gurmukhi digits and achieved recognition accuracy of 95% and 92.6% respectively.

Ubeeka Jain et al. [8] recognized Gurmukhi character set including Gurmukhi numerals, using an approach based on neocognitron and achieved the recognition accuracy of 92.78%.

Kartar Singh Siddharth, et al. [9] have proposed recognition of handwritten Gurmukhi numeral using three different feature sets. They used the features based on (first) distant profiling having 128 features, (second) projection histogram having 190 features, and (third) zonal density and background directional distribution (BDD) forming 144 features. They achieved recognition accuracy of 98%, 99.2%, and 99.13% respectively using the dataset of 1500 Gurmukhi numerals.

In the handwritten character recognition field, significant research efforts have been made. However, no machine with a performance level similar to that of humans is developed. The performance of character recognition system depends significantly upon features used. Therefore, a wide variety of approaches have been purposed to try to extract the distinctive features of handwritten characters [10]. We believe that by using the features similar to those extracted by the human visual system, the machine can achieve the performance similar to that of humans.

Gabor filters [11] [12] have been used extensively in image processing, texture analysis for their excellent properties: optimum joint spatial/spatial-frequency localization and ability to simulate the respective fields of simple cells in the visual cortex. These characteristics suggest that the Gabor filter based features seem to be similar to features extracted by humans and, thus, may be effective in classifying characters. In the middle 1990s, some researchers [12] applied Gabor filers to feature extraction for character recognition.

This paper will introduce a new and robust feature extraction method based on Gabor filters for character recognition. The main purpose of this paper is to describe the key technologies used to construct our Gurmukhi numeral recognizer. Among them, we highlight three key techniques contributing to the high recognition accuracy, namely (1) the use of Gabor features to characterize the original numeral image, (2) the use of discriminative feature extraction methods to reduce the dimensions of the feature vector and, (3) the use of SVM classifier to achieve high performance.

The features used for this work are obtained from the directional information of the image points of the numerals. A Support Vector Machine (SVM) classifier has been used for the recognition of off-line handwritten numeral of Gurmukhi Language. In this work, at first, the bounding box of a numeral is segmented into blocks and directional features are computed in each of these blocks. Next, these blocks are down sampled by a Gaussian filter. Finally, the features obtained from the down sampled blocks are fed to the classifier for recognition.

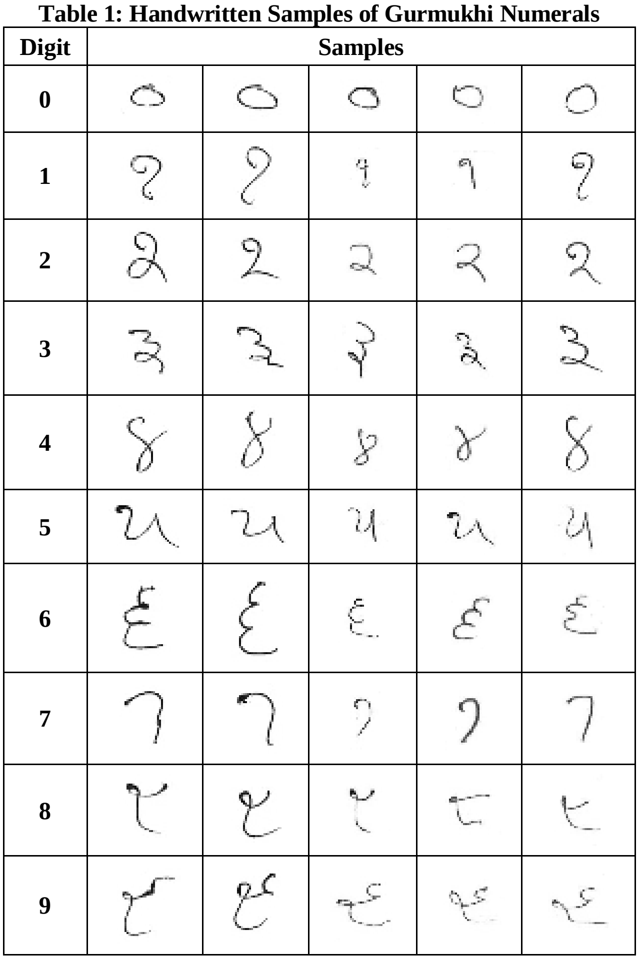
This paper is organized as follows. In section 2, we describe a Gurmukhi numeral dataset used for supporting this work. In section 3, the details of the proposed Gabor filter based feature extraction for robust numeral recognition are given. Then the SVM classifier used to validate our work is given in section 4. In section 5, the obtained results are listed.

2. DATASET

The dataset of Handwritten Gurmukhi numerals used to implementation this work is same as that used in [9]. This dataset is collected from 15 different persons, each of which writing 10 samples for each 10 Gurmukhi Numerals, on a white paper in an isolated manner.

The table 1 shows some samples of the collected dataset used to implement the work of this paper. Among these samples, some distortions and irregularities are also incorporated by writers.

Table 1: Handwritten Samples of Gurmukhi Numerals



3. FEATURES EXRACTION

Feature extraction is an integral part of any recognition system. The aim of feature extraction is to describe the pattern by means of minimum number of features that are effective in discriminating pattern classes. We have used following sets of features extracted to recognize Gurmukhi characters.

1. Gabor Features – GABM
2. Gabor Features – GABN

3.1 Gabor Feature Extraction

Gabor filters are defined by harmonic functions modulated by a Gaussian distribution. The use of the 2D Gabor filter in computer vision was introduced by Daugman in the late 1980s. Since that time it has been used in many computer vision applications including image compression, edge detection, texture analysis, object recognition and facial recognition.

Marcelja and Daugman discovered that simple cells in the visual cortex can be modelled by Gabor functions [13]. The 2D Gabor functions proposed by Daugman are local spatial bandpass filters that achieve the theoretical limit for conjoint resolution of information in the 2D spatial and 2D Fourier domains.

Families of self-similar 2D Gabor wavelets have been proposed and adopted for image analysis, representation, and compression (e.g., [14]). Gabor filters have also been used extensively in various computer vision applications such as texture analysis, texture segmentation and classification, edge detection, etc. Furthermore, features extracted by using Gabor filters (we call them Gabor features) have been successfully applied to many pattern recognition applications such as face recognition, iris pattern recognition, and fingerprint recognition. It is interesting to notice that in OCR area Gabor features have not become as popular as they have in face and iris pattern recognition areas. This situation is difficult for the new comers to understand, especially considering the following facts:

1) Gabor features are well motivated and mathematically well-defined,

2) They are easy to understand, fine-tune and implement,

3) They have also been found less sensitive to noises, small range of translation, rotation, and scaling.

3.1.1 Introduction to Gabor Filter

Gabor filters have been used extensively in image processing, texture analysis for their excellent properties: frequency and orientation representation of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination.

A Gabor Filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function.

h (x, y) = g(x, y) s(x, y)

Where s(x,y) is a complex sinusoid, known as carrier and g(x,y) is a Gaussian shaped function, known as envelope. The Gabor filters are self similar, i.e. all filters can be generated from one mother wavelet by dilation and rotation. Thus the 2-D Gabor filter with the response in spatial domain is given by Eq. (1) and in spatial-frequency domain is given by Eq.(2).

Since Gaussian Function is a complex function so on convolving Gabor Filter with input image the output obtained can be used in various ways. Two of ways of manipulating the output of Gabor Filter to extract features are described below.

$$h(x, y; \lambda, \phi, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\frac{{R_1}^2}{\sigma_x^2} + \frac{{R_2}^2}{\sigma_y^2}\right]\right\} \times \exp\left[i \cdot \frac{2\pi R_1}{\lambda}\right] \qquad (1)$$

where

$$R_1 = x\cos\phi + y\sin\phi,$$
$$R_2 = -x\sin\phi + y\cos\phi.$$

$$h(u, v; \lambda, \phi, \sigma_x, \sigma_y) = C\exp\left\{-2\pi^2\left(\sigma_x^2\left(F_1 - \frac{1}{\lambda}\right)^2 + \sigma_y^2(F_2)^2\right)\right\}, \qquad (2)$$

where

$$F_1 = u\cos\phi + c\sin\phi\ ,$$
$$F_2 = -u\sin\phi + v\cos\phi.$$
$$C = \text{const.}$$

The other form of 2-D Gabor Filter in terms of frequency can be represented as:

$$h_{x,y,\theta,f} = e^{-\frac{1}{2}\left(\frac{x'^2}{\sigma_x^2}+\frac{y'^2}{\sigma_y^2}\right)} \cdot e^{i2\pi f x} \qquad (3)$$

Where σ_x and σ_y explain the spatial spread and are the standard deviations of the Gaussian envelope along x and y directions. x' and y' are the x and y co-ordinates in the rotated rectangular co-ordinate system given as:

$$x' = x\cos\theta + y\sin\theta$$
$$y' = y\cos\theta - x\sin\theta$$

Any combination of θ and f, involves two filters, one corresponding to sine function and other corresponding to cosine function in exponential term in Eq. (3). The cosine filter, also known as the real part of the filter function, is an even-symmetric filter and acts like a low pass filter, while the sine part being odd-symmetric acts like a high pass filter.

Gabor filters having Spatial frequency (f = 0.0625, 0.125, 0.25, 0.5, 1.0) and orientation (θ =nπ/6) where n varies in the range 0 to 6, have been used in reported work.

3.2 Gabor Features-GABM

This set of features is based on extracting features from Energy magnitudes of output of Gabor Filters.

In this the output of Gabor Features is divided into 3 parts.

One part corresponds to the Real part (Re) of the Output,

Other one corresponds to the Imaginary (Im) part of output,

The last one corresponds to Absolute ($\sqrt{(Re^2 + Im^2)}$) of Complex Output of the Gaussian Gabor Filter.

After obtaining the required three forms of output, Energy magnitudes of these outputs are calculated. Energy magnitude of any output is nothing but square of that output.

In the proposed system, multi-bank Gabor filters having five different values for Spatial frequency (f = 0.0625, 0.125, 0.25, 0.5, 1.0) and seven different values for orientation θ = (0, 30, 60, 90,120, 150, 180) are chosen thus giving total of 35 combinations of Gabor filters. From the output of each Gabor filter Real, Imaginary and Absolute part of output are calculated and then for each part mean (μ) and standard deviation (σ) are computed, which serves as Gabor features. Thus for each character image we get a feature vector of 210 values given by

$$F = \begin{bmatrix} \mu_{(re)_1}, \sigma_{(re)_1}, \mu_{(im)_1}, \sigma_{(im)_1}, \mu_{(abs)_1}, \sigma_{(abs)_1}, \dots\dots \\ \dots\dots\ \mu_{(re)_{70}}, \sigma_{(re)_{70}}, \mu_{(im)_{70}}, \sigma_{(im)_{70}}, \mu_{(abs)_{70}}, \sigma_{(abs)_{70}} \end{bmatrix}$$

3.3 Gabor Features-GABN

This set of features is based on extracting features from real parts and imaginary parts of output of Gabor Filters.

In this also the output of Gabor Features is divided into 2 parts, Real part and Imaginary part. For this set of features we don't process the outputs further as we did in earlier technique rather we use the outputs as it is, as our feature extracted. One thing to note is that whenever the Image is convolved with Gabor Filter the size of output is similar to size of input image we have taken. Since size of image being 32x32 the output of convolution is also 32x32 thus making the feature extracted with dimensionality of 1024. The processing time and storage increases proportionally with increase in dimensionality of feature vector. Since the size of feature is very high, the required processing time and storage can be reduced by the dimension reduction employing the principal component analysis (PCA transform). The principal component analysis is a typical dimension reduction procedure based on the orthonormal transformation which maximizes the total variances, and minimizes the mean square error due to the dimension reduction. It is shown that the dimensionality can be reduced to 1/5 without sacrificing the recognition accuracy. Thus by applying PCA we have reduced the dimensionality of feature vector from 1024 to 200.

For this set of features we have to determine the optimum combination of θ & f out of the above mentioned ranges of θ and f. Along with varying values of both θ & f we also need to determine right pair of values of (σ_x, σ_y) to obtain the most suitable result as feature extracted. For our approach $\sigma_x = 7$, $\sigma_y = 6$, θ = pi/6, f = 0.05 serves as the optimum set of values.

4. CLASSIFICATION

Support Vector Machines (SVM) classifier

Support vector machines (SVM) are a group of supervised learning methods that can be applied to classification or regression. The standard SVM classifier takes the set of input data and predicts to classify them in one of the only two distinct classes. SVM classifier is trained by a given set of training data and a model is prepared to classify test data based upon this model. For multiclass classification problem, we decompose multiclass problem into multiple binary class problems, and we design suitable combined multiple binary SVM classifiers. Our problem also needs to classify the characters into 10 different classes of Gurmukhi numerals. We obtained such multiclass SVM classifier tool LIBSVM available at [15]. A practical guide for SVM and its implementation is available at [16].

According to how all the samples can be classified in different classes with appropriate margin, different types of kernel in SVM classifier are used. Commonly used kernels are: Linear kernel, Polynomial kernel, Gaussian Radial Basis Function (RBF) and Sigmoid (hyperbolic tangent).

The effectiveness of SVM depends on kernel used, kernel parameters and soft margin or penalty parameter C. The common choice is RBF kernel, which has a single parameter gamma (g or γ). We also have selected RBF kernel for our experiment. Radial Basis Function (RBF) kernel, denoted as

$$K(x_i, x_j) \equiv \exp(-\gamma \parallel x_i - x_j \parallel^c)$$

Best combination of C and γ for optimal result is obtained by grid search by exponentially growing sequence of C and γ and each combination is cross validated and parameters in Combination giving highest cross validation accuracy is selected as optimal.

In N-fold cross validation we first divide the training set into N equal subsets. Then one subset is used to test by classifier trained by other remaining N-1 subsets. By cross validation each sample of train data is predicted and it gives the percentage of correctly recognized dataset.

5. RESULT AND ANALYSIS

7-fold Cross Validation

In our implementation we have used 7-fold cross validation. First we created randomly generated 7-fold cross-validation index of the length of size of dataset. This index contains equal proportions of the integers 1 to 7. These integers are used to define a partition of whole dataset into 7 disjoint subsets. We used one division for testing and remaining divisions for training. We did so 7 times, each time changing the testing dataset to different division and considering remaining divisions for training. Thus we got 7 sets of feature vectors containing training and testing dataset in the size ratio of 6:1.

The average recognition accuracy of these randomly generated 7 sets of training and testing is referred as cross validation accuracy. For selection of these parameters to obtain optimized results, first we used small sample of whole dataset and observed the parameters giving highest results. Later we refined this optimization by cross validation of whole dataset.

In SVM classifier, the results vary significantly on small values of C. These results are more sensitive to change with parameter gamma (g) of RBF kernel comparative to C. At larger values of C results are stable and variation is negligible. Most of the results of SVM listed are observed at larger range of C tested up to 500, while the values of kernel parameter used vary from (0.01 to 2). As the value is increased beyond this range accuracy decreases gradually.

The recognition rate for different values of gamma (γ) at C =4 is as shown in figure 2.

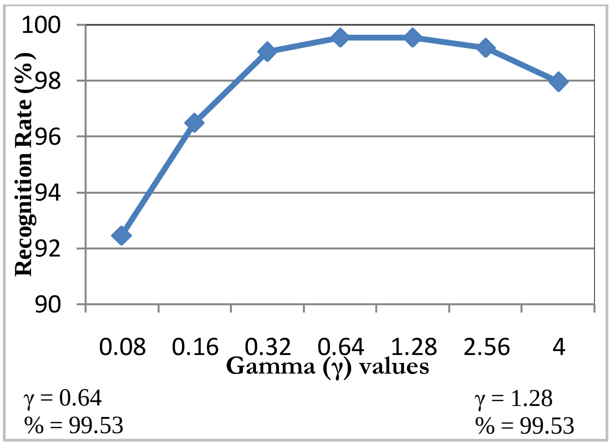


Figure 2: Graph between Recognition Rate and Gamma

The table 2 depicts the optimized results obtained with different features set at refined parameters.

Table 2: Parameters Used For Feature Set

Feature Set	Recognition Rate	Parameters
Gabor Feature GABM(210)	98.4%	C=4; γ = 256
Gabor Feature GABN(200)	99.53%	C=4; γ = 0.64-1.28

Table 3 illustrates past work done in the recognition of Gurmukhi Numerals and comparison of work done in this paper with them.

Table 3: Comparison of accuracy with different methods

S. No.	Method	Accuracy (%)
1	Structural and Statistical Features	92.6
2	Neocognitron Approach	92.78
3	Projection Histogram	99.13
4	Zonal Density and BBD	99.2
5	Purposed System	99.53

While observing the results at other values of parameter C, it is analyzed that decreasing the value of C irrespective of any change in γ slightly decreases the recognition rate, but on increasing the value of C and after a certain increment normally after 64 i.e. at higher values of C the recognition rate becomes stable. In contrast, the recognition rate always changes with the change in γ.

6. CONCLUSION AND FUTURE SCOPE

Thus we can conclude that we have obtained the maximum recognition rate as 99.53% by using GABN one of variant of Gabor Filter output as a Feature Extractor of dimensionality 200. The purpose of using Gabor Filters as mode of feature extractor is to promote its utility as major feature extraction technique in field of character recognition of Indian Scripts especially Gurmukhi. Very less literature is available on utilization of Gabor Filters for character Recognition. The work can be extended to increase the results by using or adding some more relevant features along with Gabor features. We can determine optimum combinations of standard deviation along x (σ_x), standard deviation along y (σ_y), orientation (θ), and frequency (f) which would yield higher recognition accuracies. We can use some features specific to the mostly confusing numerals, to increase the recognition rate. We can divide the entire data set to apply specific and relevant features differently. More advanced classifiers as MQDF or MIL can be used and multiple classifiers can be combined to get better results.

7. REFERENCES

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

[2] Bhowmick, T. K., et al., Proc.9th Int. Conf. on Information Technology, pp. 105-110, 2006. An HMM based recognition scheme for handwritten Oriya numerals.

[3] Sharma, N., Pal, U., and Kimura, F., Proc.9th Int. Conf. On Information Technology, pp. 133-136, 2006. Recognition of handwritten Kannada numerals.

[4] Bajaj, R., Dey, L. and Chaudhury, S., Sadhana, vol.27, pp.-59-72, 2002. Devnagari numeral recognition by combining decision of multiple connectionist classifiers.

[5] Bhattacharya, U. et al., In Proc. 10th IWFHR, pp.613-618, 2006. Neural combination of ANN and HMM for handwritten Devnagari Numeral Recognition.

[6] Jangid, M., Singh, K., Dhir, R., and Rani, R., Internation Journal of Computer Applications (IJCA), Vol. 22, No.1, May 2011. Performance Comparison of Devanagari Handwritten Numerals Recognition.

[7] Sharma, D., Lehal,G. S., Preety Kathuria, MORC Spain, 2009. Digit Extraction and Recognition from Machine Printed Gurmukhi Documents.

[8] Ubeeka Jain, Sharma, D., International Journal of Computer Applications (IJCA), Vol. 4, No. 8, 2010. Recognition of Isolated Handwritten Characters of Gurumukhi Script using Neocognitron.

[9] Siddharth, K. K., Dhir, R., and Rani, R., International Journal of Computer Application (IJCA), Vol. 28, No. 2, August 2011. Handwritten Gurmukhi Numeral Recognition using Different Feature Sets.

[10] Suen, C.Y., Signal Process. 4, 193-207(1982). Distinctive features in the automatic recognition of handwritten characters.

[11] Daugman, J. G., J. Opt. Soc. Am. A, Vol. 2, No. 7, July 1981.1. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters.

[12] Hamamoto, Y., Uchimura, S., Pattern Recognition, Vol 19, pp. 45-52, 2001. A Gabor filter-based method for recognizing handwritten numbers.

[13] Gabor, D., J. IEE, Vo1.93, pp.429-459, 1946. Theory of communication.

[14] Daugman, J. G., IEEE Trans. onASSP, Vo1.36, No.7, pp.1169-1179, 1988. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression.

[15] Chang, C. C., and Lin, C. J., 2001. LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[16] Hsu, C. W., Chang, C. C., and Lin, C. J., A Practical Guide to Support Vector Classification", [Online]. Available at : http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.