Perform Evaluation of Various Feature Extraction Techniques for Gurmukhi Script

Pardeep Kaur
Dept. of Information Technology
Chandigarh University
Gharuan, Mohali, Punjab, India

Shamandeep Singh
Assistant Professor
Dept. of Computer Science & Engineering
Chandigarh University
Gharuan, Mohali, Punjab, India

ABSTRACT

Optical character recognition (OCR) is more popular system for research field. In India, a lot of work has been done on various scripts, like Devanagari, Bangla, Tamil, Gujarati, Oriya etc. In this paper, we present a review of various OCR systems for Gurmukhi scripts. All the feature extraction techniques for extracting the features, classifiers for classification are discussed in various section of this paper. At the end, we have concluded various techniques on the basis of their accuracy.

Keywords
OCR, HCR, Gurmukhi script.

1. INTRODUCTION

At the present, we are being influenced a bunch of computers and all the processing is done automatically. According to today's improvement, it is more important that the transfer of data between computers and humans should be simple and fast. Recognition play a bigger role in data transfer between computers and humans [1]. Automatic recognition of handwritten and printed characters present on documents like envelopes, forms, cheques, post offices, commercial application in banks, libraries, and publishing house [2].

Optical Character Recognition, usually abbreviated as OCR [3], is the process of converting scanned images or hand-written text into a machine editable text such as ASCII code [2]. Several optical character recognition (OCR) systems are offered commercially within the market [4]. OCR is wide used to convert books and documents into electronic files, to change record-keeping in associate workplace, to publish the text on an internet site, 3D object recognition, or to automatic cartography etc [4] [5]. Nowaday, OCR is one of the most challenging area of research [1].

Optical Character Recognition deals with the difficulty of recognizing optically processed characters [6]. Basically, OCR classified into two types as off-line and on-line. In off-line recognition, the character has been written initial and recognition has been performed afterward [7] [8]. But, in on-line system, data are captured throughout the writing procedure with the help of a special electronic pen [3].

Gurmukhi script is the world's 14th most widely spoken language [5]. Several other Indian languages like Gujarati, Hindi, and Bangali scripts are similar to Gurmukhi [2]. In short, the advancements in OCR system of Indian scripts including Tamil, Oriya, Gujarati, Gurmukhi, Kannada, Bangla, Telugu and Devanagari up to 2012 can be seen in [2] [4]. In this paper, we present a review of all the improvements till 2014 in handwritten Gurmukhi script recognition along with their performance. The comparison is completed with regard to feature set, classifier, and according accuracy rate.

We also give a comparison of all the reported methods in tabular form. We also summarize the various techniques being applied for feature extraction and classification.

This paper is organized as follows, Section 2 describe the basic properties of Gurmukhi script. Section 3 describes the various feature extraction techniques for Gurmukhi script. Section 4 compares all the various feature extraction techniques. Conclusions are provided in section 5.

2. BASIC PROPERTIES OF GURMUKHI SCRIPT

In India, there are 18 official languages namely Assamese, Bangla, Gujarati, Gurmukhi, Hindi, Kannada, Kashmiri, Konkani, English, Malayalam, Marathi, Rajasthani, Nepali, Oriya, Sanskrit, Sindhi, Tamil, Telugu, and Urdu [9]. Gurmukhi is the official language of Punjab and is spoken by 27 million people in the world [10]. The name Gurmukhi derived from the Punjabi term "Guramukhi", which means "from the mouth of the Guru" [1]. In this section, we describe the various features of Gurmukhi script. For designing a OCR system, these specific features play an important role [4].

    i.   Gurmukhi script is cursive in nature and the character set consist of 3 vowel bearers, 32 consonants, 6 additional consonants, 9 vowel modifiers, 3 auxiliary signs and 3 half characters [5] [3]. The character set of Gurmukhi script is shown in Figure 1.

The consonants

ਸ ਹ ਕ ਖ ਗ ਘ ਙ ਚ ਛ ਜ ਝ ਞ ਟ ਠ ਡ ਢ ਣ ਤ ਥ ਦ ਧ ਨ ਪ
ਫ ਬ ਭ ਮ ਯ ਰ ਲ ਵ ੜ

The vowel bearers

ੳ ਅ ੲ

Additional consonants

ਸ਼ ਜ਼ ਖ਼ ਢ਼ ਗ਼ .ਲ

The Vowel Modifiers

ਾ ਿ ੀ ੈ ਿ ੀ ੋ ੁ ੂ

Auxiliary Signs

ੱ ੰ ੰ

The Half Characters

੍ਹ ੍ਰ ੍ਵ

Figure 1: Gurmukhi script character set [3]

    ii.   Writing style is from left to right and top to bottom [11].

iii.   Gurmukhi script is written below the line and there is no concept of upper and lower case characters [*12*].

iv.   Most of the Gurmukhi characters have a horizontal line at the upper part called headline [*1*] and characters of word are connected by this line [*13*].

v.   A text in Gurmukhi script can be divided into three horizontal zones namely upper zones, middle zone and lower zone [*11*], as shown in Figure 2. The upper zone represents the area above the headline where half characters, vowel modifiers and other symobols are appear [*12*] while the middle zone denotes the area below the headline where the consonants and some sub-parts of vowels are present [*13*]. The lower zone denotes the region below middle zone where half characters and some vowels are lie in the foot of consonants [*5*].
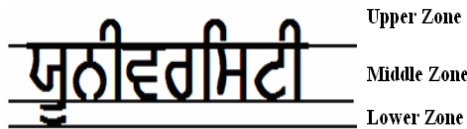
Figure 2: Horizontal zones [*5*]

vi.   In Gurmukhi script, there are lots of similar character pairs. Some similar pairs are: ਖ and ਸ਼, ਵ and ੲ, ੜ and ਡ, ਸ and ਸ਼, ਪ and ਪ

3. PRELIMINARY WORK ON FEATURE EXTRACTION TECHNIQUES

In this section, we introduce various feature extraction techniques for Gurmukhi script. This section is divided into 2 parts: (i) handwritten character (ii) handwritten numeral. In this survey, we only review the offline handwritten characters/numerals recognition.

3.1.Handwritten character recognition

The major difficulty in handwritten character recognition is the variability of writing styles between two different writers. In this paper, they proposed two feature extraction techniques namely parabola fitting and power curve fitting techniques. They also analyzed the performance of other techniques like zoning, diagonal, directional, gradient and chain code features. The classifiers that are used namely Support vector machine and k-NN classifiers with 3 flavors, i.e., SVM with linear kernel, SVM with polynomial kernel and SVM with RBF kernel. The system achieves a recognition accuracy of 98.10 and 97.14% using k-NN and SVM classifiers [*1*] [*8*].

They have been presented a paper on offline handwritten Gurmukhi character recognition system. They proposed two feature extraction techniques namely diagonal features and transition features with k-NN classifier. When they have used the diagonal features with k-NN classifier, the maximum recognition accuracy achieved 94.12%. The recognition accuracy can be improved by considering a large dataset of characters for training the classifiers [*3*]. In this paper, they has been used a 48*48 pixels normalized image and created 64 (8*8) zones and used zoning densities of these zones as features. The values in feature vector are normalized in the range 0 to 1. They observed 73.02% highest accuracy with SVM kernel with Poly. Kernel [*5*].

They presented a paper of Gurmukhi character recognition of isolated handwritten characters by using Neocognitron

technique. Neocognitron is used as a feature extractor and classifier. In this paper they found the recognition accuracy of both learned and unlearned images of characters. The overall accuracy for both learned and unlearned characters are 92.78% [*11*].

In this paper, they presented a system for recognition of machine printed Gurmukhi script. Recognition of Gurmukhi script faces the lots of problems mainly unique characteristics of script like connectivity of characters, similar characters. In segmentation part, a word image breaks into 3 zones and the character image is further breaks into sub-character or sub-symbol image. A multi-stage classifier is used to classify the characters. The recognition rate achieved by this system is 96.6% with 175 c/s speed [*13*].

In this paper, they presented a novel hierarchical technique for isolated offline handwritten Gurmukhi character recognition. They proposed four types of topological features, namely, horizontally peak extent features, vertically peak extent features, diagonal features, and centroid features. For classification support vector machine (SVM) has been used. SVMs classifier has been considered with four different kernels, namely, polynomial kernel, RBF kernel, linear kernel, and sigmoid kernel. For training and testing, they have been used 3,500 samples of character images. The highest accuracy achieved by proposed technique is 91.80% [*14*].

They presented a paper in which they deal with weka based classification methods for offline handwritten Gurmukhi character recognition. In this they have been used two efficient feature extraction techniques, namely, parabola curve fitting and power curve fitting based features. For recognition, they have been used 3500 samples of isolated handwritten Gurmukhi characters. They presented a novel framework for Gurmukhi character recognition using weka based classifiers and provides new standard for future investigation. The maximum recognition accuracy achieved by parabola curve fitting based features and multilayer perceptron model classifier are 82.92%. In this work, they have been used c language and weka classification software tool [*15*].

They presented a paper in which a grading system for Punjabi writers based on offline handwritten Gurmukhi characters recognition. They proposed four feature extraction methods, namely, zoning, diagonal, directional, intersection and open end points and Zernike moments feature. For classification, k-NN, HMM and Bayesian classifiers are used. They also compare the handwriting of one writer with other writers. This approach can also be extended for other Indian scripts such as Bengali, Tamil and Devanagri [*16*].

In this paper, they used 2 techniques for extracting the features namely, parabola curve and power curve fitting based feature. The classifier that has been used in this work is SVM with three flavors, namely, SVM with polynomial kernel, SVM with linear kernel and SVM with RBF kernel. The highest recognition accuracy of 97.14% is achieved using parabola curve fitting based features is applied on SVM classifier [*17*].

In this paper, they proposed an offline handwritten Gurmukhi character recognition system. The diagonal features and intersection and open end points features methods are used for extracting the features. The classifier that has been used in this work is SVM with three flavors: and SVM with RBF kernel, SVM with polynomial kernel & SVM with linear kernel. The maximum recognition accuracy of 94.29% is

achieved. This accuracy can be increased by considering a larger data set while training the classifier [18].

They presented a paper in which they have been used some statistical features like zonal density, projection histograms, and distance profiles. In addition, they also used background directional distribution (BDD) features for extracting the features. 200 samples of Gurmukhi characters are used for training and testing. For classification, they have been used k-NN, SVM and PNN. The maximum accuracy obtained is 95.04% using zonal density and BDD feature methods with SVM classifier [19].

In this paper, the main focuses on the extraction and recognition of Digits. In this paper, first extracted the Gurmukhi digits from printed documents and then recognized. The system is divided into three stages. In first, segmentation of input image and separate the different logical parts. In second, feature extraction stage analyzes the various structural and statistical features like zoning; directional distance distribution etc. In final stage, classification is done on extracting features. K-nearest neighbor classifier is used for classification purpose. The maximum recognition accuracy achieved is 92.6% [20].

In this paper, they presented an OCR for handwritten Gurmukhi characters. For extracting the features, different Daubechies Discrete Wavelet Transform techniques are used. Back propagation neural network is used for classification. An average accuracy of 93.41% is achieved by proposed method. In this paper, 400 samples of first five characters are used for training and 100 samples are used for testing the classifier [21].

In this paper, they proposed an isolated handwritten Gurmukhi character recognition system. They used Gabor filter based method for extracting the features. For dataset, 200 samples of 35 characters are collected from different persons. SVM classifier is used for classification purpose. The maximum accuracy achieved is 94.29% with purposed method [22]. Table 1 shows an overview of various feature extraction techniques & classifiers for handwritten character recognition.

Table 1: Existing feature extraction techniques & classifiers for handwritten character recognition

Authors	Input Pattern	Features	Classifiers
Kumar M et al. [1]	Handwritten characters	Parabola curve and power curve fitting features	SVM, k-NN classifiers
Kumar M et al. [3]	Handwritten characters	Diagonal & Transition features	k-NN classifier
Sharma DV & Jhajj P [5]	Handwritten characters	Zoning density	SVM & k-NN classifiers
Sharma D and Jain U [11]	Handwritten characters	Neocognitron technique	Neocognitron technique
Lehal G S and Singh C	Handwritten characters	Zoning density	Multi-stage classifier

Authors	Input Pattern	Features	Classifiers
[13]			
Kumar M. et al. [14]	Handwritten characters	Horizontally peak extent features, vertically peak extent features, diagonal features, and centroid features.	SVM classifier
Kumar M. et al. [15]	Handwritten characters	Parabola curve and power curve fitting features	Multilayer perceptron model
Kumar, M et al. [16]	Handwritten characters	Zoning, diagonal, directional, intersection and open end points and Zernike moments feature.	k-NN, HMM and Bayesian classifiers
Jindal, M.K et al. [17]	Handwritten characters	Parabola curve and power curve fitting features	SVM with RBF kernel
Kumar M et al. [18]	Handwritten characters	Diagonal features and intersection and open end points features	SVM classifier
Siddharth K.S et al. [19]	Handwritten characters	zonal density, projection histograms, and distance profiles & BDD features	k-NN, SVM and PNN classifiers
Sharma, D. V., et al. [20]	Handwritten Digits	structural and statistical features	K-NN
Singh P and Budhiraja S [21]	Handwritten characters	Daubechies Discrete Wavelet Transform technique	Back propagation Neural network
Singh. S., et al. [22]	Handwritten Characters	Gabor Filter	SVM with RBF Kernel

3.2. Handwritten Numeral recognition

In this paper, Wavelet Transform method is used for feature extraction. Zonal densities of different zones of numeral images have been used in feature set. For classification, Back propagation neural network is used. A maximum recognition accuracy achieved by this method is 88.83% [23].

In this paper, they presented three feature sets and three classifiers for handwritten Gurmukhi numerals. The first feature set consists of distance profiles having 128 features. The second feature set consists of different types of projection histograms having 190 features. Third feature set consists of zonal density and Background Directional Distribution (BDD)

having 144 features. The SVM, PNN and KNN classifiers are used for classification. The SVM classifier is used with RBF kernel. With KNN & PNN, the maximum accuracy achieved by third feature set is 98.33% & 98.51% while with SVM the maximum accuracy achieved by second feature set is 99.2%. SVM classifiers achieved the highest accuracy as compared with PNN and KNN [24].

In this paper, they have used two feature set techniques that are GABM and GABN for recognition which are based on Gabor filter. For classification, SVM classifier with RBF kernel is used. The maximum accuracy is 99.53% achieved with second feature set and 98.4% with first feature set [25].In this paper, they present zone based hybrid approach for extracting the features which is the combination of zone centroid zone and image centroid zone. They obtained 99.73% recognition accuracy [26]. Table 2 shows an overview of various feature extraction techniques & classifiers for handwritten numerals recognition.

Table 2: Existing feature extraction techniques & classifiers for handwritten numerals recognition

Authors	Input Pattern	Features	Classifiers
Singh P & Budhiraja S [23]	Handwritten Numerals	Wavelet transform feature	Back propagation neural network
Siddharth, K.S et al. [24].	Handwritten Numerals	distance profiles, projection histograms, zonal density and BDD	SVM, PNN & KNN classifiers
Singh, S., & Dhir, R. [25]	Handwritten Numerals	GABM and GABN based on Gabor Filter	SVM with RBF kernel
Rani. et al. [26]	Handwritten Numerals	Zone based hybrid approach	SVM classifier

4. COMPARISON OF FEATURE EXTRACTION TECHNIQUES

In this section, we compare the various feature extraction techniques. We have organized the comparison into 2 set of tables as given below.

Table 3: For Handwritten character

Table 4: For Handwritten numeral

We also give the summery of different feature sets and classifiers used by Gurmukhi` scripts.

4.1 Feature Sets

The performance of OCR system depends on the feature set extracted from the character image. We have observed that 8 different feature sets have been used for OCR system. These are as follows.

I. Diagonal features [3]: Diagonal features are used to achieve higher accuracy and reducing misclassification.

II. Transition features [16]: The Transition technique works on the raw character image and does not require resizing. It is used to determine character confidences in our initial segmentation system.

III. Projection Histogram features [19]: It counts the number of pixels in specified directions. Three types of projection histogram are created namely, horizontal, vertical, left diagonal & right diagonal.

IV. Wavelet Transform [27]: The main advantage of wavelet filters is that they provide a graceful tool for multi resolution analysis. It decomposes the input image into hierarchy of several levels of resolution. Other important advantage of using wavelets is to reduce improve accuracy, ambiguity, fast computations and practical applications.

V. Zoning densities [19]: The main goal of zoning is to obtain the local characteristics instead of global characteristics. In zoning, each image is divided into N×M zones.

VI. Intersection features [4]: It is defined as a pixel point which has more than two neighboring pixels in 8-connectivity while open end has exactly one neighbor pixel. Intersection points are unique for a character in different segment.

VII. Chain code feature [4]: The chain code for a character contour will give a smooth, unbroken curve as it grows along the perimeter of the character & completely cover the character.

VIII. Zernike moments [16]: Zernike introduced a set of complex polynomials { $Vnm(x, y)$ } which forms a complete orthogonal set over a unit disk $x^2 + y^2 \leq 1$. The form of the polynomial is:

$$Vnm(x, y) = Vnm(p, \theta) = Rnm(p).exp(jm\theta)$$

Where $p = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1} y/x$ and $j = \sqrt{-1}$

IX. Intersection and open end points [16]: An intersection point is the pixel that has more than one pixel in its neighborhood and an open end point is the pixel that has one pixel in its neighborhood.

X. Statistical features [4]: These are helpful to degraded documents where reliable extraction of topological or structural features becomes difficult due to losing of some black pixels. Statistical features perform well across various fonts. Table 3-4 shows the comparison between various feature sets and classifiers.

Table 3: comparison of Gurmukhi scripts for handwritten characters

Authors	Features	classifiers	Dataset	Training	Testing	Accuracy (%)
Kumar M et al.(2014)	Parabola curve and power curve fitting features	SVM, k-NN classifiers	3500	3465	35	98.10%
Kumar M et al. (2011)	Diagonal & Transition features	k-NN classifier	3500	1750	1750	94.12%
Sharma DV & Jhajj P (2010)	Zoning density	SVM & k-NN classifiers	5125	3075	2050	73.02%
Lehal G S and Singh C (2000)	Zoning density	Multi-stage classifier	------	------	-------	96.6%
Kumar M. et al. (2014)	Hierarchical Technique	SVM classifier	3500	------	-------	91.80%
Kumar M. et al. (2014)	Parabola curve and power curve fitting features	Multilayer perceptron model	3500	2100	1400	82.92%
Kumar, M et al. (2011)	Zoning, diagonal, directional, intersection and open end points and Zernike moments feature.	k-NN, HMM and Bayesian classifiers	3500	------	------	------
Jindal, M.K et al. (2011)	Parabola curve and power curve fitting features	SVM with RBF kernel	3500	------	------	97.14%
Kumar M et al.	Diagonal features and intersection and open end points features	SVM classifier	3500	3150	350	94.29%
Siddharth K.S et al. (2011)	zonal density, projection histograms, and distance profiles & BDD features	k-NN, SVM and PNN classifiers	200	-----	------	95.04%
Sharma D and Jain U (2010)	Neocognitron technique	Neocognitron technique	300	150	150	92.78%
Singh P and Budhiraja S (2012)	Daubechies Discrete Wavelet Transform technique	Back propagation Neural network	500	400	100	93.41%
Sharma, D. V.,et al. (2009)	structural and statistical features	K-NN	-------	-------	-------	92.6%
Singh. S., et al. (2012)	Gabor Filter	SVM with RBF kernel	200	------	------	94.29%

Table 4: comparison of Gurmukhi scripts for handwritten numerals

Authors	Features	classifiers	Dataset	Training	Testing	Accuracy (%)
Singh P & Budhiraja S (2012)	Wavelet transform feature	Back propagation neural network	1000	750	250	88.83%
Siddharth, K.S et al. (2012)	distance profiles, projection histograms, zonal density and BDD	SVM, PNN & KNN classifiers	1500	------	------	99.2%
Singh, S., & Dhir, R. (2012).	GABM and GABN based on Gabor Filter	SVM with RBF kernel	1500	------	-------	99.53% with GABN % 98.4% with GABM
Rani. et al. (2012)	Zone based hybrid approach	SVM with RBF classifier	1500	1050	450	99.73%

4.2 Classifiers

Like feature sets, classifiers play an important role in OCR systems. Various classifiers have been used for Gurmukhi script. Some are explain here:

I.   Support vector machine (SVM): It constructs a hyper plane or set of hyper planes in a high or infinite dimensional space, which can be used for classification, regression, or other tasks [28].

II. K-nearest neighbor: It is a non-parametric method used for classification and regression. It is a type of instance based learning or lazy learning. It is the simplest algorithm in machine learning algorithms [29].

III. Neural network: It consists of units (neurons), arranged in layers, which convert an input vector into some output. Each unit takes an input, applies a function to it and then passes the output on to the next layer [30]. The multi layer perceptron classifier has been very popular in offline character recognition [4].

IV. Naive Bayes classifiers: These are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem [31].

5. CONCLUSION

In this paper, we have reported the various work done on Gurmukhi script. We have organized the review around work have done on handwritten characters/numerals. In this review paper, compared the various feature extraction techniques, classifiers & different datasets which are used in Gurmukhi script for improving the recognition performance.

6. REFERENCES

[1] M, sharma, R.K & Jindal, M.K kumar, "Efficent feature extraction techniques for offline handwritten Gurmukhi character recognition," National Academy Science Letters, vol. 37(4), pp. 381-391, 2014.

[2] R., Kolhe, S.R., Patil, P. M., & Pal, U. Jayadevan, "Offline recognition of Devanagari script: A survey.," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 41(6), pp. 782-796, 2011.

[3] M., Jindal, M. K., & Sharma, R. K. Kumar, "K-nearest neighbor based offline handwritten Gurmukhi character recognition.," In Image Information Processing (ICIIP), 2011 International Conference on IEEE, pp. 1-4, November 2011.

[4] S., & Harit, G. Bag, "A survey on optical character recognition for Bangla and Devanagari scripts.," Sadhana, vol. 38(1), pp. 133-168, 2014.

[5] D., & Jhajj, P. Sharma, "Recognition of isolated handwritten characters in Gurmukhi script," International Journal of Computer Applications, vol. 4(8), pp. 9-17, 2010.

[6] L. Eikvil, "Optical character recognition," citeseer. ist. psu. edu/142042. html, 1993.

[7] V. P. Agnihotri, "Offline Handwritten Devanagari Script Recognition," International Journal of Information Technology and Computer Science (IJITCS), vol. 4(8), p. 37, 2012.

[8] M., Sharma, R. K., & Jindal, M. K. Kumar, "Offline handwritten gurmukhi character recognition: Study of

[9] U., & Chaudhuri, B. B. Pal, "Indian script character recognition: a survey," pattern Recognition, vol. 37(9), pp. 1887-1899, 2004.

[10] A. M. M., & Dhanya, P. M. Chacko, "Handwritten Character Recognition In Malayalam Scripts-A Review," arXiv preprint arXiv:1402.2188, 2014.

[11] D., & Jain, U. Sharma, "Recognition of Isolated Handwritten Characters of Gurumukhi Script using Neocognitron," International Journal of Computer Applications, vol. 10(8), pp. 10-16., 2010.

[12] K. S., Dhir, R., & Rani, R. Siddharth, "Handwritten Gurumukhi Character Recognition Using Zoning Density and Background Directional Distribution Features," International Journal of Computer Science and Information Technologies, vol. 2(3), pp. 1036-1041., 2011.

[13] G. S., & Singh, C. Lehal, "A Gurmukhi script recognition system," in InPattern Recognition, 2000. Proceedings. 15th International Conference on IEEE, 2000, pp. 557-560.

[14] M., Jindal, M. K., & Sharma, R. K. Kumar, "A Novel Hierarchical Technique for Offline Handwritten Gurumukhi Character Recognition," National Academy Science Letters, vol. 37(6), pp. 567-572, 2014.

[15] M., Jindal, M. K., & Sharma, R. K. Kumar, "Weka-Based Classification Techniques for Offline Handwritten Gurumukhi Character Recognition," in In Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012, Springer India., 2014, pp. 711-720.

[16] M., Jindal, M. K., & Sharma, R. K. Kumar, "Classification of characters and grading writers in offline handwritten Gurumukhi script," in In Image Information Processing (ICIIP), 2011 International Conference on, IEEE, 2011, pp. 1-4.

[17] M., Jindal, M. K., & Sharma, R. K. Kumar, "Offline Handwritten Gurumukhi Character Recognition Using Curvature Feature," in In Proc. of International Conference on AMOC, 2011, pp. 981-989.

[18] M., Sharma, R. K., & Jindal, M. K. Kumar, "SVM Based Offline Handwritten Gurumukhi Character Recognition," pp. 1-12.

[19] K. S., Jangid, M., Dhir, R., & Rani, R. Siddharth, "Handwritten Gurumukhi Character Recognition Using Statistical and Background Directional Distribution," 2011.

[20] D. V., Lehal, G. S., & Kathuria, P. Sharma, "Digit extraction and recognition from machine printed Gurumukhi documents," In Proceedings of the International Workshop on Multilingual OCR, ACM, p. 12, July 2009.

[21] P., & Budhiraja, S. Singh, "OCR for Handwritten Gurumukhi Script using Daubechies Wavelet Transforms," International Journal of Computer Applications, vol. 45(10), pp. 25-28, 2012.

[22] Aggarwal. A., & Dhir. R. Singh. S., "Use of Gabor Filters for Recognition of Handwritten Gurmukhi Character," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 5(2), pp. 234-240, 2012.

[23] P., & Budhiraja, S. Singh, "Offline Handwritten Gurmukhi Numeral Recognition using Wavelet Transforms," International Journal of Modern Education and Computer Science (IJMECS), vol. 4(8), p. 34, 2012.

[24] K. S., Dhir, R., Rani, R., Jangid, M., Dhir, R., Rani, R., & Singh, K. Siddharth, "Comparative Recognition of Handwritten Gurmukhi Numerals Using Different Feature Sets and Classifiers," In Proceedings of International Conference on Image Information Processing (ICIIP 2011), 2011.

[25] S., & Dhir, R. Singh, "Recognition of Handwritten Gurmukhi Numeral using Gabor Filters.," International Journal of Computer Applications, vol. 47(1), pp. 7-11, 2012.

[26] G. S. R., & Dhir, R. Rani, "Handwritten Gurmukhi Numeral Recognition using Zone-based Hybrid Feature Extraction Techniques.," 2012.

[27] Navghane, A., & Dandawate, Y. Dixit, A., "Handwritten Devanagari character recognition using wavelet based feature extraction and classification scheme," in In India Conference (INDICON), 2014 Annual IEEE, 2014, pp. 1-4.

[28] http://en.wikipedia.org/wiki/Support_vector_machine.

[29] http://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm

[30] http://www.robots.ox.ac.uk/~dclaus/digits/neural.htm

[31] http://en.wikipedia.org/wiki/Naive_Bayes_classifier