IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 6, Ver. V (Nov.-Dec. 2016), PP 41-49
www.iosrjournals.org

Serial Number Recognition in Banknotes Using HoG Feature Extraction and KNN Classification

Er.Ramanjit Kaur[1], Er. Priyadarshni[2]
[1](Department of ECE), Student, Ludhiana College of Engineering and Technology, India
[2](Department of ECE), Associate Professor, Ludhiana College of Engineering and Technology, India

Abstract: *Character recognition is one of the many complex areas of pattern recognition with immense research and practical importance. This paper presents serial number recognition in Indian banknotes which can help to manage the currency circulation, financial transactions and prevents forgery crimes by detecting fraudulent banknotes. This approach consists of various steps including preprocessing, feature extraction and classification. In preprocessing, linear contrast stretching and morphological opening is performed for the gray level normalization and removal of irrelevant pixels from the degraded images. A well known Otsu method is applied on images containing bimodal histogram pattern resulting binary images for further analysis. Histogram of oriented gradients (HoG) feature extraction method has been used to extract feature vectors from the concatenated histograms containing intensity gradients and orientation data which are computed from the pixels within 4 by 4 cell size. These feature vectors are processed by K-nearest neighbor (KNN) classifier for the character recognition task. Experimental results show that the recognition rate of 83.81% is achieved by employing the proposed binarization methods performing with the precision rate of 90% and recall rate of 78%.*
Keywords: *Character recognition, linear contrast stretching, HoG feature extraction, KNN classification.*

I. Introduction

Handwritten and printed character recognition is an area of pattern recognition which defines an ability of a machine to analyze different patterns and identify the unique character. It is of great importance in the field of financial transactions including bank cheque processing [1], currency recognition [2] and other banking services. This paper focuses on the investigation of the serial numbers in Indian banknotes which can help to detect the fraudulent banknotes and manage the currency circulation.

In the recent years, a few papers have been published for serial number recognition based on genetic algorithm artificial neural network [3] and support vector machines (SVM) [4]. A part based method for the recognition of RMB (renminbi bank note, the paper currency used in China) serial numbers is presented in which a set of local image parts are extracted from the training samples using DoG keypoint detector. Then, the SVM classifier is trained and tested with these parts. Each part of the sample is recognized by the classifier and the final category is determined by combining the classification results of all these parts using strategies called max rule, major voting and multiple voting [5]. For handwritten digit recognition [6], a trainable feature extractor based on LeNet 5 convolutional neural network and SVM classifier has been employed and tested on the widely known MNIST database performing with the recognition rate of 99.78%. A serial number recognition system [7] is proposed by comparing recently emerged techniques in character recognition including different types of feature extraction methods, classifiers, multiple classifier combination strategies, distortion methods and rejection schemes and conducting evaluations on the NUST-RMB 2013 database introducing a noval cascade rejection method which gives 100% reliability by rejecting only 1.01% of the test samples. For the recognition of paper currency from different countries, a texture based feature extraction method using Hidden Markov Model has achieved 98% accuracy [8].

The identification of serial numbers in banknotes is a challenging process due to poor image quality caused by uneven illumination, background intensity variation, inaccuracy in character extraction and so on. Thus, taking these challenges under consideration a serial number recognition system is designed in several stages. All these stages have been discussed in the following sections. Section 2 explains the preprocessing operations including linear contrast stretching and thresholding techniques applied on input samples. The HoG feature extraction method has been employed to extract the global features as described in Section 3. Section 4 discusses the recognition task and lists all the experimental results being evaluated. Finally, the conclusions are drawn in Section 5.

II. Preprocessing

The experiments for different techniques related to preprocessing, feature extraction and classification are performed on serial number database which has been collected from daily used banknotes. Fig.1 shows a scanned banknote image in which unique serial numbers composed of 9 alphanumeric characters are printed at the lower left and the upper right corners respectively. The first step in the character extraction process is to detect the serial number region from the prior knowledge of serial numbers' size and location [9]. The region of interest (ROI) in the image is outlined using bounding box as shown in Fig. 2. The ROI images with uneven illumination are then operated by the preprocessing methods to minimize gray-scale variances and to eliminate noises in the background.

Fig. 1. Scanned banknote image.

Fig. 2. Serial numbers detected in banknote using bounding box.

A large number of preprocessing techniques have been introduced in the previous papers which have reported excellent performances in their respective applications. For car license plate binarization [10], the Otsu's algorithm has been an effective method but it is not compatible with serial number binarization due to the complex texture, uneven lumination and intensity variations. In addition, the thresholding techniques based on adaptive image contrast [11] and 8 bit plane slicing [12] are presented for document image binarization which need complex calculation work. The proposed approach is simple and effective on degraded images. It applies image enhancement methods namely morphological opening and linear contrast stretching separately on images with minimum contrast and uneven illumination. Further, the binarization is performed by choosing a suitable threshold value.

2.1 Morphological Operations:-

The serial number images in Fig. 3a and Fig3b exhibit non-uniform background. Thus, the first step is to estimate a background approximation from the original image using morphological opening to overcome the intensity variations. The opening operation makes use of two processes, erosion (removing pixels from the image) and dilation (adding pixels to the image) with the same structuring element creating the background approximation image by removing the foreground pixels which are completely covered under the disk shaped structuring element of radius equal to 15. Then, the resulting image is subtracted from the original image producing another image with the uniform background as depicted in Fig. 3c and Fig. 3d.

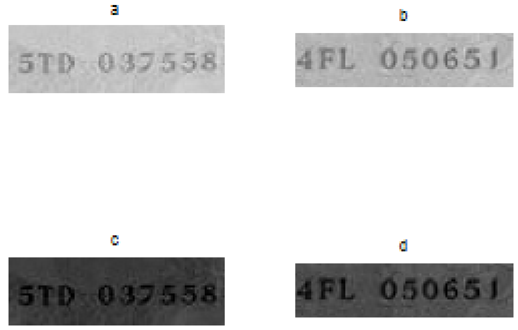

Fig. 3. Preprocessed serial number images: Original images: (a) and (b), Images with uniform background: (c) and (d).

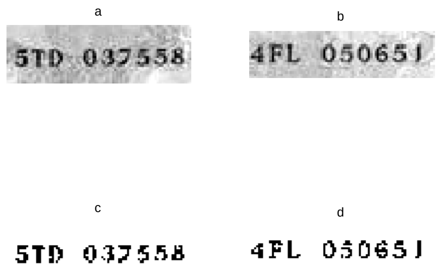

Fig. 4. Contrast adjusted images: (a) and (b), binary images: (c) and (d).

Further, the contrast adjusted images (Fig. 4a and Fig. 4b) are converted to binary images by choosing a suitable threshold value that can segment the serial numbers from the background accurately as shown in Fig. 4c and Fig. 4d.

2.2 Linear Contrast Stretching:-

The gray scale normalization of the images with uneven illumination is performed through linear contrast stretching. This method increases the contrast level by transforming an image I with intensity values in the range (min,…..,max) into a new image I_N with intensity values in the range (newMin,…..,newMax). Fig. 5a shows an image of a character sample with the intensity range from 50 to 140 approximately which is predicted from its histogram (Fig. 6a). In its normalization process, these values are linearly stretched to the new intensity values in the range (0, 255) by using formula [13] given in equation (1).

$$I_N = (I - Min) \times \frac{newMax - newMin}{Max - Min} + newMin \qquad (1)$$

The normalized image presented in Fig. 5b along with its stretched histogram (Fig. 6b) is then binarized by the computed threshold value, as illustrated in Fig. 5c.

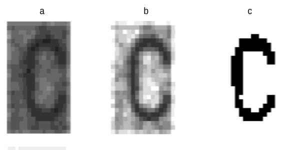

Fig. 5. Linear contrast stretching: (a) original image, (b) contrast adjusted image, (c) binary image.

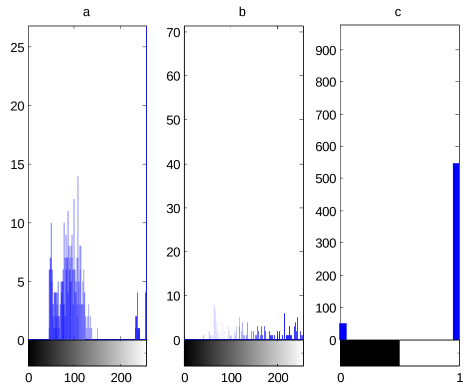

Fig. 6. Histograms for: (a) original, (b) contrast adjusted and (c) binary images.

2.3 Otsu's Algorithm for Binarization:-

In the related study, the synthetic images of characters (0 to 9 digits and A to Z alphabets) are being used to train the classifier. These images contain bimodal histogram pattern with two classes of pixels (foreground and background pixels). Thus, Otsu's method is efficient for synthetic samples and provides effective binarization results as depicted in Fig. 7. It computes a global threshold value separating the background and foreground such that their intra-class variance is minimal while the inter-class variance is maximal [14]. The intra-class variance is the variance within the class i.e. either from background or foreground. It is defined as a weighted sum of variances of the two classes as:

$$\sigma_w^2(t) = w_0(t)\,\sigma_0^2(t) + w_1(t)\,\sigma_1^2(t) \tag{2}$$

$$where$$

$$\sigma_0^2(t) = E\left[(x-\mu_0)^2\right]$$

$$\sigma_1^2(t) = E\left[(x-\mu_1)^2\right]$$

Where w and σ denotes the probabilities of the pixels and variances of two classes respectively, t is the threshold separating these two classes. E is the expected/average value of the set of pixels in each class, μ_0 and μ_1 corresponds to the mean values of both the classes. The inter-class variance is the variance of one class compared with the other class as defined below:

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) \tag{3}$$

$$= w_0(t)[\mu_0(t) - \mu_T(T)]^2 + w_1(t)[\mu_1(t) - \mu_T(t)]^2 \tag{4}$$

$$= w_0(t)w_1(t)[\mu_0(t) - \mu_1(t)]^2 \tag{5}$$

The global threshold value (t) should maximizes this inter-class variance, $\sigma^2_b(t)$ as much as possible to obtain satisfied binary results. The separation of pixels in two classes using threshold value is illustrated in Fig. 8.

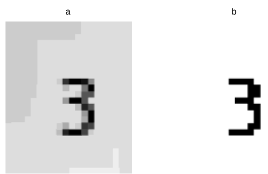

Fig. 7. Training samples: (a) gray-scale image, (b) binary image.

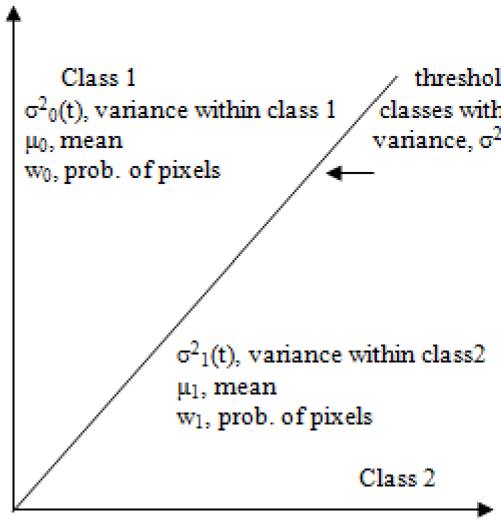

Fig. 8. Separation of pixels in two classes using global threshold value.

III. Hog Feature Extraction

Feature extraction can also be thought of as a preprocessing step that reduces the dimension of data by extracting relevant information. There are numerous features associated with character recognition e.g. zoning feature, structural feature, directional feature, crossing points and contour feature. A good set of features play an important role in the recognition process by retrieving information used to identify the unique characters. Structural features that provide the local information about the object have been utilized for the recognition of Oriya and Gujarati scripts [15] and [16] respectively. A trainable feature extractor (TFE) based on LeNet5 convolutional neural network extracts the topological features in its first layers and classify them with its last layers [6]. Ref. [17] employs the scale invariant feature transform (SIFT) algorithm which extracts local feature points of the Chinese characters from all the possible angles.

The proposed work makes use of gradient directional features which have been extracted using the histogram of oriented gradients (HoG) descriptor [18] due to the reason that it encodes the shape information of the object to be detected by the distribution of intensity gradients. In this method, an image (16×16) containing a digit is divided into small connected rectangular regions called R-HOG blocks. Each block contains a group of 2×2 cells whose size is selected according to the number of features required to provide the fine details of the object. These blocks overlap by 50% and can be determined by three parameters i.e. number of cells per block, number of pixels per cell and number of histogram channels. In this study, a 4 by 4 cell size is preferred which results in the creation of 8 by 8 blocks. For the pixels within each cell, a histogram of gradient directions is calculated by evaluating the gradient and orientation computations.

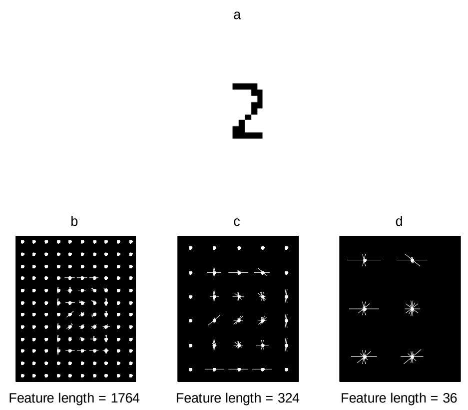

Fig. 9. HoG feature extraction of character sample with different cell sizes: (a) Binary image (b) Cell size = [2 2] (c) Cell size = [4 4] (d) Cell size = [8 8].

3.1 Gradient computation:-

For the implementation of this algorithm, the first step is to calculate the gradient values [13] of an image as follow:

$$g_x = \frac{\partial f}{\partial x} \quad (6)$$

$$g_y = \frac{\partial f}{\partial x} \quad (7)$$

∂f/∂x and ∂f/∂y are the gradients in x and y directions respectively. These values can be calculated by convolving 1-dimensional filters with the image. These filters are the row and column vectors having value [-1 0 +1].

3.2 Orientation binning:-

The second step is to calculate the orientation values with formula given in equation (8) using the results obtained in gradient computation:

$$\theta = \tan^{-1}\left[\frac{g_y}{g_x}\right] \quad (8)$$

Both these calculations are used to create the cell histograms with 9 orientation based channels [18] which are evenly spread over 0 to 180 degree.

In the final step, the HoG descriptor concatenates the cell histograms from all the block regions from which the feature vectors are created and returned in a 1 by N vector, where N is the HoG feature length. The visualization function is used to show the extracted features within a cell. The size of the cell is specified in pixels as a 2-element vector. With the variation in cell size parameter, the amount of shape information encoded in the feature vector also varies. For example, to capture large scale spatial information, the cell size needs to be increased while this results in losing the small scale detail. The sub figures in Fig. 9 are depicting that with the cell size of [8 8], the HoG feature vector of 36 dimensions is obtained, which does not provide much shape information. While the cell size of [2 2] provide a lot of information but increases the feature length up to 1764. This is the primer reason to select a 4 by 4 cell size which extracts the feature vectors with 324 dimensions encoding enough spatial information to identify the shape of character samples and helps in speed up the training of classifier. The feature lengths for different cell sizes are calculated as described below:

Feature length for [8 8] cell size = 1 × 1 (block) × 4 (number of cells) × 9 (bins) = 36
Feature length for [4 4] cell size = 3 × 3 (blocks) × 4 × 9 = 324
Feature length for [2 2] cell size = 7 × 7 (blocks) × 4 × 9 = 1764

Fig.10 shows the flow chart of HoG feature extraction and KNN classification based character recognition system. The recognition procedure has three main steps: First and foremost is the preprocessing step which is performed both on training and testing samples respectively. Then, the HoG descriptor is deployed to obtain directional features from the characters, last but not the least is training and testing of the KNN classifier with the extracted features to achieve the desirable recognition results.

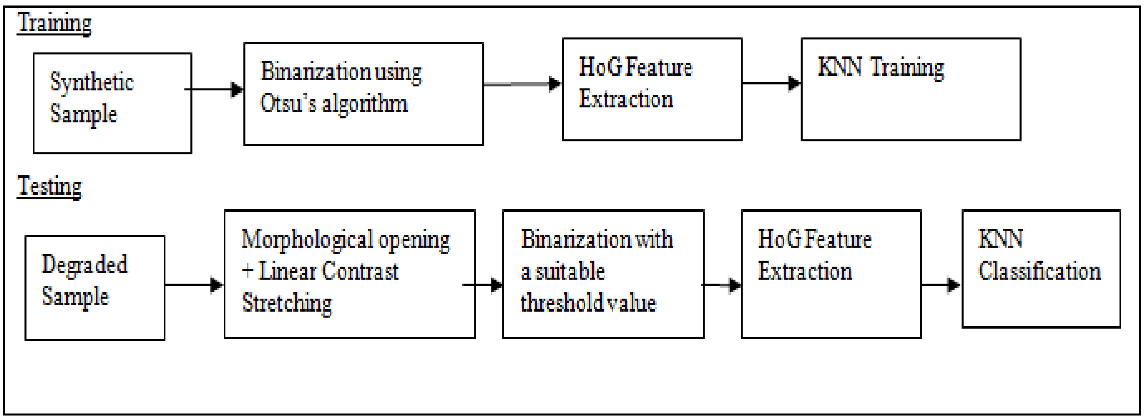

Fig. 10. Working flow diagram of the proposed serial number recognition system.

IV. KNN Classifier

The recognition task is performed through nearest neighbor (KNN) classifier which is a non-parametric statistical classifier [19]. It is trained through supervised learning algorithm in which the labeled training data is used to create a predictor model. This model predicts the category of new samples by assigning them to the classes among its k-nearest neighbors. A common weighting scheme is used to allot each neighbor a weight of 1/d, where d is the distance to the neighbor from the query point. The weight value determines the contribution of neighbors in the classification process i.e. the neighbor with the largest weight value contributes more by assigning its class because it is nearer to the query point as compare to the distant ones. The distance to the new point from each neighbor is calculated using Euclidean distance as:

$$d_{st}^2 = (y_s - y_t)(y_s - y_t)' \tag{9}$$

y_s and y_t are the training vectors in 2-dimensional space. With the default value of K = 1, the class of the closest training sample is assigned to the new sample.

The data (HoG features), processed by the classifier is obtained from 180 images containing alphanumeric samples of 35 categories i.e. there are 5 images per sample. These images are used in training and the samples collected from bank note images are utilized for the testing purpose. The predictor model contains HoG feature values which are specified as matrix of scalar values (n × 324) where n is the number of samples/observations and there are 324 feature vectors corresponding to each sample. The classification values are specified as a cell array of strings whose each row represents the category of the corresponding observations. This model uses the predict method which assigns class labels (Y) to the new observations (X new) with the minimal expected cost and maximal posterior probability. It returns a vector of predicted class labels for X new which is defined as follow:

$$y = \sum_{k=1}^{K} P\left(k/x\right) C\left(y/k\right) \tag{10}$$

where y denotes the predicted labels, K is the number of classes, P(k/x) is the posterior probability of class k for observation x and C(y/k) is the cost of classifying as y when its true class is k.

The posterior probability or conditional probability of class k for observation x can be defined as the product of its prior probability, P(x/k) with the likelihoods P(k) and P(x) as given below:

$$P\left(k/x\right) = \frac{P\left(x/k\right) \times P(k)}{P(x)} \tag{11}$$

The two costs related with KNN classification are: true misclassification cost, C(i,j) and expected misclassification cost, C(n,k). C(i,j) is the cost of assigning an observation to class j if its true class is i. If i ≠ j then C(i,j) = 1 which corresponds to incorrect classification and if i = j then C(i,j) = 0 which corresponds to correct classification. C(n,k) is the expected or average cost of assigning n observations to each of the k classes and is defined as:

$$C(n,k) = \sum_{i=1}^{K} P\left(i/Xnew(n)\right) C\left(k/i\right) \tag{12}$$

where K is the number of classes, P(i/X new(n)) represents the posterior probabilities of classes (i= 1 to K) for observations X new (1 to n) and C(k/i) is the true misclassification cost of assigning an observation to k when its true class is i.

A good recognition system is known for its accuracy which is calculated using formula:

$$\text{Re}\,cognition\,Rate = \frac{Number\,of\ correctly\,recognized\ samples}{Total\,number\,of\ samples} \times 100 \tag{13}$$

The experiments of classification for digits and alphabets are performed separately using MatLab software [20]. The recognition rate of 80% is achieved for digits classification with the expected cost of 12% which states incorrect classification for two digits, 6 and 8 while digit 5 is misclassified. Similarly, for alphabet classification, a recognition rate of 92.8% is achieved with few false positive rates.

4.1 Result Analysis:-

The performance of the proposed binarization methods can be determined from the precision and recall rates. The precision rate is the fraction of retrieved information to the relevant information which measures how much relevant information is retrieved or extracted while recall rate is the fraction of retrieved relevant information to the total relevant information that would have been retrieved. Another parameter known as F-measure or F-score is used to measure the accuracy of the algorithms from the harmonic mean of precision and recall rates as below:

$$F - score = 2 \times \frac{\Pr ecision \times \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \qquad (14)$$

Table 1 shows the outputs of few measures used to determine the performance of binarization techniques discussed in section 2.

Table 1: Binarization Results

Method	Precision Rate	Recall Rate	F-Measure
Otsu's Method for synthetic images.	92%	60%	72.63%
Proposed method (combined with morphological opening and linear contrast stretching) applied on test images.	90%	78%	83.57%

Table 2 shows the recognition results obtained after implementing the proposed algorithms.

Table 2: Recognition Results

Character Samples	Recognition Rate
Digits	81.25%
Alphabets	92.8%
Average Recognition Rate	83.81%

V. Conclusions

Serial number recognition system based on HoG feature extraction method and KNN classifier is successfully implemented using impressive binarization techniques for preprocessing. A database of alphanumeric samples is collected from synthetic images and serial numbers in banknotes. Linear contrast stretching and morphological opening is applied to adjust the contrast level of images and to minimize the intensity variations in the background which have reported the precision and recall rates of 90% and 78% respectively with the F-measure of 83.57%. The HoG feature extraction method is deployed to extract the gradient directional features providing essential shape information of the character samples. At the end, the KNN classifier uses extracted HoG feature vectors for classification producing results with the average recognition rate of 83.81% and classification cost of 12% with some misclassifications leading to the further study on serial number recognition with more representative data set containing hundreds of character samples using different feature extraction methods and classifiers to produce better recognition results.

References

[1] Qizhi Xu, Louisa Lam, Ching Y. Suen: 'Automatic segmentation and recognition system for handwritten dates on Canadian bank cheques'. Proceedings of the 7th International Conference on Document Analysis and Recognition, Edinburgh, UK, August 2003, pp. 704-708.

[2] R. Vishnu, Bini Omman: 'Principal component analysis on Indian currency recognition'. Proceedings of the 5th International Conference on Computer and Communication Technology, Allahabad, India, September 2014, pp. 291-296.

[3] Ting-Ting Zhao, Ji-yin Zhao, Rui-rui Zhang, Lu-lu Zhang: 'Study on RMB number recognition based on genetic algorithm artificial neural network'. Proceedings of the 3rd International Congress on Image and Signal Processing, Yantai, China, October 2010, Volume 4, pp. 1951-1955.

[4] Wenhong Li, Wenjuan Tian, Xiyan Cao, Zhen Gao: 'Application of support vector machine (SVM) on serial number identification of RMB'. Proceedings of the 8th World Congress on Intelligent Control and Automation, Jinan, China, July 2010, pp. 6262-6266.

[5] Bo-Yuan Feng, Mingwu Ren, Xu-Yao Zhang, Ching Y. Suen: 'Part-based high accuracy recognition of serial numbers in bank notes', N.El Gayar et al. (Eds.): ANNPR, LNAI 8774, Springer, Switzerland, 2014, pp. 204-215.

[6] Fabien Lauer, Ching Y. Suen, Gérard Bloch: 'A trainable feature extractor for handwritten digit recognition', Pattern Recognition 40, 2007, pp. 1816-1824.

[7] Bo-Yuan Feng, Mingwu Ren, Xu-Yao Zhang, Ching Y. Suen: 'Automatic recognition of serial numbers in bank notes', Pattern recognition 47, 2014, pp. 2621-2634.

[8] Hamid Hassanpour, Payam Masoumi Farahabadi: 'Using Hidden Markov Models for paper currency recognition', Expert Systems with Applications 36, 2009, pp. 10105-10111.

[9] Bo-Yuan Feng, Mingwu Ren, Xu-Yao Zhang, Ching Y. Suen: 'Extraction of serial numbers on bank notes'. Proceedings of the 12th International Conference on Document Analysis and Recognition, Washington, USA, August 2013, pp. 698-702.

[10] Zhang Zhi-Yong, Song Yang: 'License plate image binarization based on Otsu algorithm and MatLab realize'. Proceedings of the International Conference on Industrial Control and Electronics Engineering, Xian, China, August 2012, pp. 1657-1659.

[11] Bolan Su, Shijian Lu, Chew Lim Tan: Robust document image binarization technique for degraded document images, IEEE Transactions on Image Processing, April 2013, Volume 22, No. 4, pp. 1408-1417.

[12] Karthika M, Ajay James: 'A proposed method for document image binarization based on bit plane slicing'. Proceedings of the International Conference on Advances in Engineering and Technology Research, Unnao, India, August 2014, pp. 1-5.

[13] Rafael Gonzalez, Richard Woods: 'Digital Image Processing', 2nd edition, Prentice Hall Upper Saddle River, New Jersey 07458, 2002.

[14] Nobuyuki Otsu: 'A threshold selection method from gray-level histograms', IEEE Transactions on Systems, Man and Cybernetics, 1979, Volume 9, No. 1, pp. 62-66.

[15] Aditya Raj: 'An optical character recognition of machine printed Oriya script'. Proceedings of the 3rd International Conference on Image Information Processing, Waknaghat, India, December 2015, pp. 543-547.

[16] Hetal Thaker, Chandresh Kumbharana: 'Structural feature extraction to recognize some of the offline isolated handwritten Gujarati characters using decision tree classifier', International Journal of Computer Applications, 2014, Volume 99, No. 15, pp. 0975-8887.

[17] Wang Zhen-yan: 'Chinese character recognition method based on image processing and Hidden Markov Model'. Proceedings of the 5th International Conference on Intelligent Systems Design and Engineering Applications, Hunan, China, June 2014, pp. 276-279.

[18] Navneet Dalal, Bill Triggs: 'Histograms of oriented gradients for human detection'. Proceedings of the IEEE Conference Computer Vision and Pattern Recognition, San Diego, CA,USA, June 2005,Volume 1, pp. 886-893.

[19] Keinosule Fukunaga: 'Introduction to statistical Pattern Recognition', 2nd edition, Academic Press, San Diego, 1990.

[20] Alasdair McAndrew: 'An introduction to Digital Image Processing with MATLAB', Australia Thomson Course Technology, 2004.