TEXT AREA IDENTIFICATION FOR RECOGNIZING DESTINATION PLACES FROM VEHICLES

Selvanayaki K.S, Somasundaram Rm,
Department of Computer Applications, Easwari Engineering College, Chennai, India.
Department of Computer Science, SNS College of Engineering, Coimbatore, India.
kssindia2004@yahoo.co.in

Abstract - Nowadays, automatic detection of text from the vehicles is an important problem in many applications. Text information present in an image can be easily understood by both human and computer. It has wide applications such as license plate reading, sign detection, identification of destination places, mobile text recognition and so on. This problem is challenging due to complex backgrounds, the non-uniform illuminations, variations of text font, size and line orientation. Once the text is identified, it can be analyzed, recognized and interpreted. Hence, there is a need for a better algorithm for detection and localization of text from vehicles. A method is proposed for detecting text from vehicles. The method makes use of features such as Histogram of oriented Gradients (HOG) and Local Binary Pattern (LBP). These features are stored which can be further used for feature matching at the time of classification. After the text region is being detected, it can be further subjected to character segmentation and recognition thereby identifying the destination places. The ability to recognize text area from the vehicles, especially buses has obvious applications like traffic management in the bus stands. The obtained results are verified and performance parameters like speed, precision and recall are determined.

Keywords- HOG, LBP, Profile based features, Skew detection and removal, Eigen value regularization

I. INTRODUCTION

Automatic detection of interest regions is an active research area in the design of machine vision systems and is used in many applications such as tourist's assistant systems, mobile robot navigation, vehicle license plate detection and recognition. Vision systems are mainly focused on constantly monitoring traffic and observe passing vehicles, extracting important features such as vehicle type, color and distinct marks.

The text content present in the vehicles is a unique feature which can be used for identifying vehicles in video surveillance applications. Text is a distinct mark that can be found in many vehicles. This paper deals with the development of an algorithm for detecting the text area from the vehicles especially buses for recognizing the places. In the bus stands, there may be circumstances where people rushing behind each and every bus to know its destination place. This situation may lead to a crowded stand, thereby increasing the chances of accidents. So if we can develop a system which will detect the destination place written on the bus boards and display the name of the place on a screen displayed in the waiting room, it will be helpful for the people as well as the policemen to control the crowd. Here, the images of buses are captured from the camera which is mounted at the entrance of the bus stands. On vehicles, text marks are very small compared to the imaged scene and are typically present on the front of the bus. There are basically three different methods for text region detection. They are 1) Texture Based 2) Connected Component Based 3) Region Based.

The texture based approach considers the text as a special texture. Here, the features are extracted over a certain region. This method makes use of the texture feature uniformity across the text regions. Then the classifier is employed to identify the existence of text region. The most commonly used method in texture based approach is to extract the features from the DCT of the text region. This is based on the observation that the text has certain horizontal and vertical frequencies. The other method employed is collecting the features of the text region from the wavelet coefficients and classify the text using SVM.

The connected component based method makes use of the observation that the text pixels are connected to each other. This method extracts the regions from the image and uses the geometric constraints to rule out non-text candidates. There are different methods for finding out the connected components. The widely used method is linking the components based on the geometric properties. The other method is finding the connected components in a troke width transformed image which is generated by shooting arrays from edge pixels along gradient direction.

The region based method utilizes the feature that the text area has distinct intensity compared with the background. The region based methods rely on the text region analysis. This method utilizes morphological operations to extract text regions. Region based methods also includes edge based methods. The edge based methods uses the observation that the text has strong edges between the character and background pixels. This

method identifies the sub structures in the image. Thus, the edges present in the image are identified and these sub structures are merged to mark the bounding boxes for text by learning based rules.

II. LITERATURE SURVEY

The proposed approach consists of mainly 5 stages – text region detection, preprocessing stage which includes skew removal and noise removal, script identification, character segmentation and character recognition. The various techniques employed for the different stages were studied and recorded.

A. Text Region Detection

AUTHOR	APPROACH	ADVANTAGES	DISADVANTAGES
Katherine L. Bouman (2011) [6]	Homogenous region selection and decomposition	Able to detect and localize both the larger texts and smaller texts	Doesnot work in blurred images
Andrej Ikica (2010) [7]	Edge profile based text detection	Able to detect the text captured from any lighting conditions	Symbols detected as text
Kai Chen (2011) [8]	Learning based method	High accuracy 95%, Less computational time	Fails when text color is similar to background
Adam Coates (2011) [9]	Unsupervised feature learning algorithm	High accuracy (85%)	Failed to detect smaller text regions
Mariano et al (2003) [2]	Color Feature Extraction	High accuracy(87%)	Unable to detect text if it is written with different colors.
Li Sun et al (2009) [4]	Corner Response Based Text Detection	It is more robust than other edge or texture based methods.	Unable to detect smaller texts

Mainly, the text detection methods can be broadly classified into three categories – Texture based, connected component based and region/edge based. Based on these differentmethods, different features are used for text detection. In the proposed approach, connected component based method is employed since each and every character is a connected component. The unwanted components can be eliminated by considering the geometric properties of the components.

B.SKEW DETECTION AND CORRECTION

AUTHOR	APPROACH	ADVANTAGES	DISADVANTAGES
Hong Liu et al (2009) [1]	**Borderline extraction using run length method**	**Efficient for complex documents**	**Accuracy may be affected by some straight lines in the non text regions**
Mehdi Felhi et al (2010) [3]	maximum gradient difference and R-signature	Able to detect multiple skews in a document	Complexity of the algorithm is more
P Shivakumara et al (2007) [10]	Boundary growing approach	Works efficiently for text binary documents	Not robust to noise
Yang Cao et al (2009) [5]	Straight line fitting	Reduces computational complexity and have better precision	Cannot detect multiple skews

The skew of the detected text region can be detected and corrected using different techniques. The most widely used methods are discussed here. The skew detection using projection profile method gives high accuracy. This method is able to detect any skew angles.

B. Script Identification

AUTHOR	APPROACH	ADVANTAGES	DISADVANTAGES
D Dhanya et al (2002) [11]	Word level identification	Accuracy is 90%	Word should have more than 5 connected components
Huanfeng Ma and David Doermann(2004)	Gabor filter analysis of textures	Accuracy rate is 90%	Single characters may not have similar texture leading to incorrect classification
Mallikarjun et al(2010) [12]	Directional visual discriminating features	Accuracy rate is 97.5%	Time complexity is more
M. C. PADMA et al (2010) [13]	Profile based features	Success rate found to be 99.5%	Training phase and testing phase required

Out of the different methods used for script identification, the proposed approach makes use of the profile based features for identifying the script. This approach mainly concentrates on 2 scripts- English and Malayalam.

D. CharacterSegmentation

AUTHOR	APPROACH	ADVANTAGES	DISADVANTAGES
Xiaodan Jia et al (2008) [14]	Vertical projection profile	Algorithm is more efficient under the condition that license plate image is degraded	Prior knowledge about the character segments are needed
P Shivakumara et al (2010) [15]	Gradient based method	Accuracy rate is 94%	Mathematical complexity is more
Huadong Xia et al (2011) [16]	Vertical projection	Works well with license plate images	Mathematical complexity is more
Youngwoo Yoon et al(2011) [17]	Blob extraction based	Accuracy rate is97.2%	Won't work in low resolution images
Youngwoo Yoon et al (2012) [18]	Lee's character segmentation algorithm	Accuracy rate is 98.3%	When input images are captured under bad lighting conditions, accuracy is less

The widely used methods for segmentation were employed. Blob extraction based method gives high accuracy. The noisy blobs can be segmented out by considering the size of the extracted blobs.

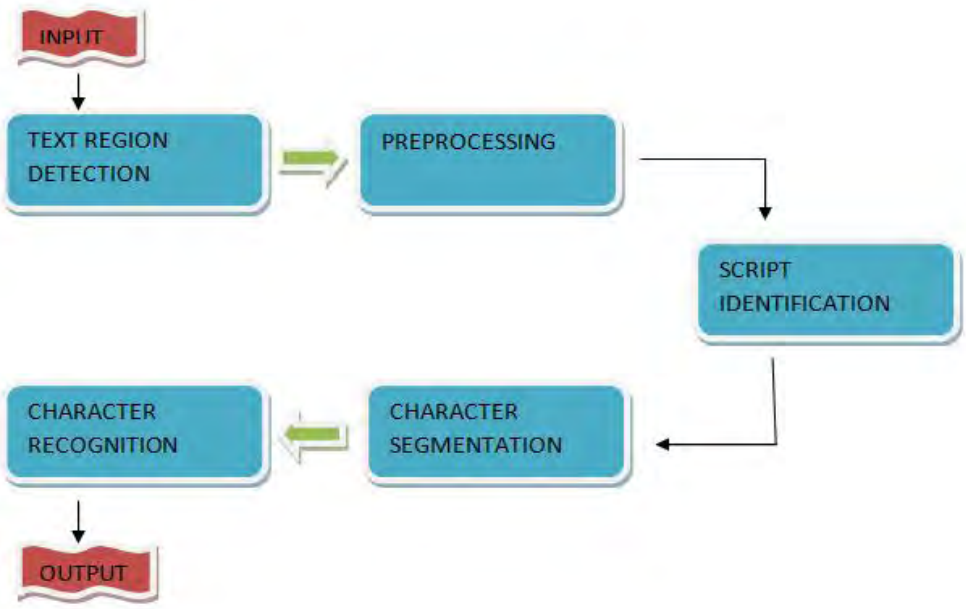
E. CharacterRecognition

Template matching was the most popular method employed for character recognition. The other approaches can be roughly classified into feature based, structural based and neural network based classification. Feature matching approach extracts different features then calculates a distance metric between test sample and the trained class. But still the selection and extraction of feature vectors remains a major issue especially if character contains noise. The paper proposes a modified approach for character recognition using Eigen feature regularization method.

AUTHOR	APPROACH	ADVANTAGES	DISADVANTAGES
J. A. Vlontzos et al (1992) [19]	Hierarchical system with Hidden Markov Model	Solves both the context sensitivity problem and the character instantiation problem	Large computational requirements - Needs O (TN2) for a N state model and T observations.
C.N. Anagnostopoulos et al (2006) [20]	Probabilistic Neural Network(PNN)	High accuracy	Design of multilayer network and computational complexity
Takayuki Kurozumi et al (1999) [21]	Eigen space method	Higher correct classification rates	May not perform well if within class variation is large
Md. Anwar Hussain et al (2010) [22]	Standard eigen space.	Lower dimensional feature vectors, $m<N$.	Low processing speed
Million Meshesha (2007) [23]	Feature extraction using PCA and LDA followed by a decision directed graph based SVM classifier	High recognition rate , Works with different fonts	Sensitive to noise, Fails in case of real life degraded documents

III.PROPOSED METHOD

OVERALL ARCHITECTURE DIAGRAM



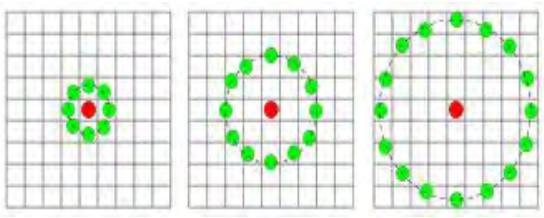
As mentioned early, the proposed approach consists of 5 stages. The text region detection phase detects the text region from the input media. This approach extracts mainly 2 features- LBP(local binary pattern) and HOG(histogram of oriented gradients). The preprocessing stage is mainly employed to remove the skew of the text region and also to remove the noise. Script identification mainly focuses on identifying the script of the document so that it can be properly fed into the appropriate character segmentation algorithm. Character recognition phase recognizes the character thereby identifying the destination place.

A. Text Region Detection

For each input image, connected component analysis is performed. The candidate regions obtained through the connected component analysis are subjected to feature extraction. The LBP and HOG features are extracted and *Local BinaryPattern (lbp)*

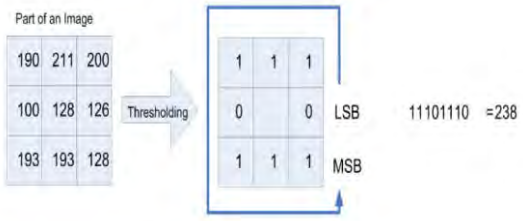
Local binary patterns (LBP) are a type of feature used for classification in computer vision. It has since been found to be a powerful feature for texture classification [24]. The LBP feature vector, in its simplest form, is created in the following manner:

- Divide the examined window to cells (e.g. 16x16 pixels for each cell).
- For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. clockwise or counter-clockwise.



Local Binary Patterns of a point P in the neighborhood region R

- Where the center pixel's value is greater than the neighbor, write "1". Otherwise, write "0". This gives an 8-digit binary number (which is usually converted to decimal for convenience).
- Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller and which are greater than the center).
- Optionally normalize the histogram.



Computing the LBP code for the central pixel, their values are compared with those stored in the database. K-NN classifier is used to classify the region into text and non-text regions.

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) * 2^p$$

$$\text{Where, } s(z) = \begin{cases} 1, & z >= 0 \\ 0, & z < 0 \end{cases}$$

- Concatenate normalized histograms of all cells. This gives the feature vector for the window.

$$h(i) = \sum_{x,y} B(LBP_{P,R}(x,y) = i) \mid i \in [0, 2^P - 1], \quad B(v) = \begin{cases} 1 & \text{when v is true} \\ 0 & \text{otherwise} \end{cases}$$

Local Binary Pattern **(**LBP*)* is a simple efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number [25]. The most

important property of the LBP operator is its robustness to monotonic gray-scale changes caused by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

Histogram of Oriented Gradients (hog)

Histogram of Oriented Gradients (HOG) is feature descriptors used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. Local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The implementation of these descriptors can be achieved by dividing the image into small connected regions, called cells, and for each cell compiling a histogram of gradient directions or edge orientations for the pixels within the cell. The combination of these histograms then represents the descriptor [27]. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination or shadowing.

The HOG feature vector, in its simplest form, is created in the following manner:

- **Gradient Computation**

The first step of calculation is the computation of the gradient values. The most common method is to simply apply the 1-D centered, point discrete derivative mask in one or both of the horizontal and vertical directions

$$G_x(x, y) = [-1\ 0\ 1] * I(x, y)$$

$$G_y(x, y) = [-1\ 0\ 1]^T * I(x, y)$$

- **Orientation Binning**

The second step of calculation involves creating the cell histograms. Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. The cells themselves can either be rectangular or radial in shape, and the histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, depending on whether the gradient is "unsigned" or "signed".

Gradient magnitude is,

$$m(x, y) = \sqrt{((d_x(x, y)^2) + (d_y(x, y)^2))}$$

Gradient angle is,

$$\theta(x, y) = \tan^{-1}(d_y(x, y) / d_x(x, y))$$

- **Descriptor Blocks**

In order to account for changes in illumination and contrast, the gradient strengths must be locally normalized, which requires grouping the cells together into larger, spatially connected blocks. The HOG descriptor is then the vector of the components of the normalized cell histograms from all of the block regions. These blocks typically overlap, meaning that each cell contributes more than once to the final descriptor.

The value of each bin is,

$$\psi_k(x, y) = \begin{cases} G(x, y) & \text{if } \theta(x,y) \in \text{bin}_k \\ 0 & \text{otherwise} \end{cases}$$

- **Block Normalization**

Blocks are L2 normalized to extract features. The final feature value is calculated as,

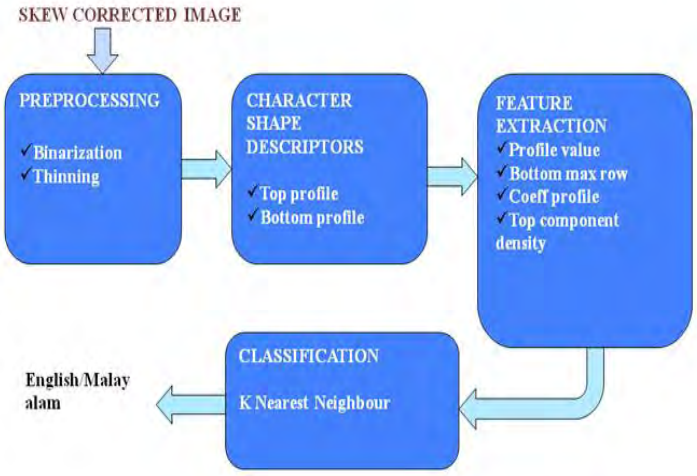
$$\sum \psi_k(x, y) + \varepsilon$$

$$f(C, B, k) = \frac{\sum_{(x,y)\in C} G(x, y)}{\sum_{(x,y)\in B} G(x, y) + \varepsilon}$$

The HOG descriptor maintains a few key advantages over other descriptor methods. Since the HOG descriptor operates on localized cells, the method upholds invariance to geometric and photometric transformations, except for object orientation. Such changes would only appear in larger spatial regions [26].

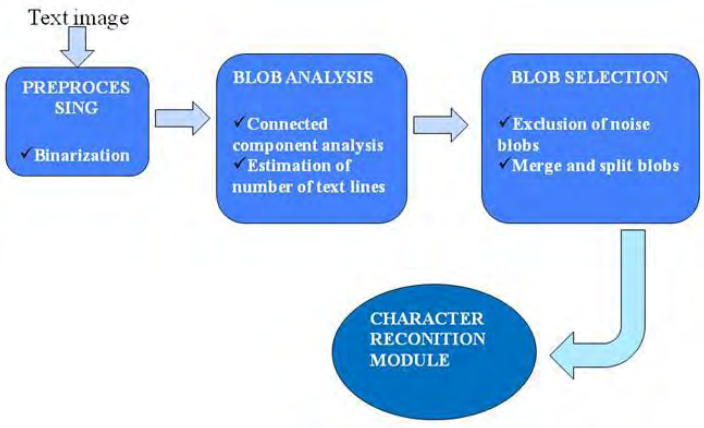
B. S*cript Identification*

The proposed model is based on the observation that every script/language has a finite set of text patterns, each having a distinct visual appearance, which helps them in recognizing the language. Every language could be identified based on its discriminating features. The proposed approach is mainly based on the concept of the top and bottom profiles of the input text lines. The character shape descriptors used in the proposed approach is top profile and bottom profile. The top profile of a text line represents a set of black pixels obtained by scanning each column of the text line from top until it reaches a first black pixel. Similarly, the bottom profile can be obtained by scanning the image from bottom to top. From the descriptors, certain features are extracted which can be efficiently used for the identification of the script. The features used in this method are Profile value, Bottom_max_row_no, Coeff_profile, Top_component_density [13].



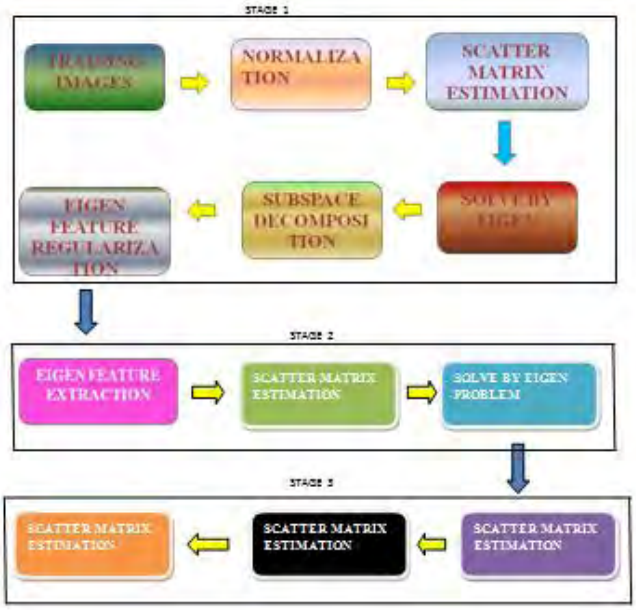
Initially the system is trained using the proposed features for different scripts. Then for the test input text image, the features are extracted and a classifier is employed to classify the text into appropriate class thereby helping in identifying the script of the document.

C. Character Segmentation



Initially, a binarization algorithm is applied to the input text image so that the foreground and background pixels can be well separated. Then, CCA is employed so that the candidate characters can be obtained. Depending upon the position of the candidate blobs, it is easy to determine the number of the text lines present in the input document. The expulsion of the noisy blobs can be done based on some geometric properties of the blobs. There may be chances that a single character may detect as multiple blobs and multiple characters may be detected as a single blob. Depending on this merging and splitting of blobs need to be done[17].

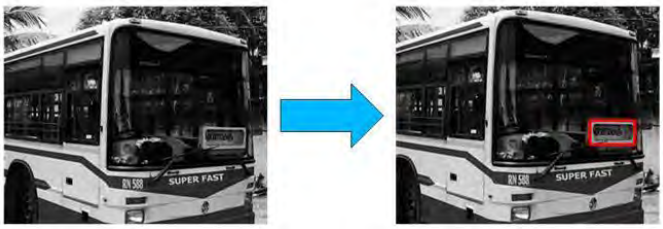
D. CharacterRecognition



This paper proposes the Eigen feature Regularization and Extraction Algorithm (ERE Algorithm) for Character Recognition [28]. This algorithm was proposed by Xudong Jiang. Here, the entire eigenspace is decomposed in to subspaces and regularization and extraction of the significant Eigen Vectors is done. The algorithm decomposes the Eigenspace spanned within class scatter matrix into face, noise and null subspaces. Eigen features are regularized differently in these three subspaces based on an Eigen Spectrum model. Then dimensionality reduction is applied thereby maximizing the variances of the extracted features and reducing the error. Finally, classifier is employed to recognize the character feature vectors.

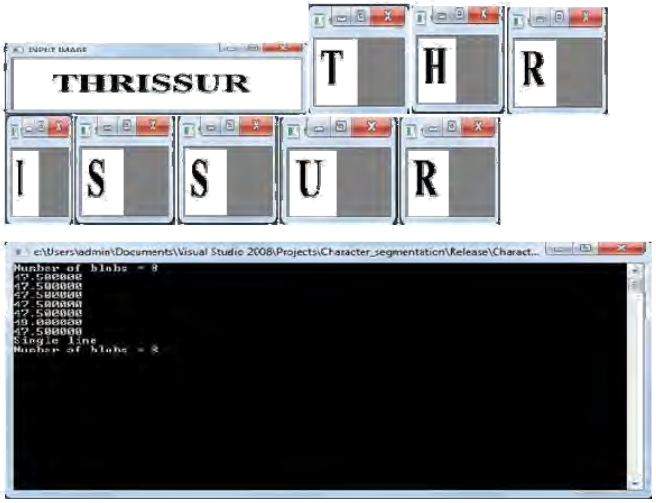
IV. EXPERIMENTAL RESULTS

The algorithm is tested with different test images and the results obtained were recorded. The result for the text region detection was found to have 87% precall rate(which is the percentage of correctly detected text area from the total detected regions). The recall value was calculated to be 76% (which is the percentage of correctly detected text area from the total regions).



In the next phase, the script of the text region need to be identified. The following output gives an idea of the top and bottom profile of the text document. From this, features discussed are extracted.

After the script is identified, the text document is fed into the appropriate segmentation algorithm. The algorithm will segment out the different characters in the text document and saves each character as an image.



The final phase is the recognition phase where the segmented characters are recognized and identified.

V. CONCLUSION

This paper proposes an approach for detecting and identifying the destination place written on the vehicles especially buses. Here, the boundary of the problem is that the destination place is written on a board placed in front of the bus. The proposed algorithm worked well and the results were analyzed and recorded. This approach mainly concentrates on the software part. Given an input image, the proposed approach detects the text region and recognizes the destination place. This work can be extended to the hardware part also. A camera can be designed so that when it is mounted at the entrance of the bus stands, it can capture the vehicles at a distance. The images or video captured from this camera can be processed using the proposed software.

ACKNOWLEDGMENT

I would like to my deep gratitude to Dr.ShivaKumar, Head, Computer Applications for enthusiastic encouragement and continuous support. I extend my thanks to Dr.Rm.Somasundaram, my research supervisor for his patient guidance and useful critiques of this work. My grateful thanks are also extended to the management, Easwari Engineering College for their continuous encouragement. Finally I wish to thank my parents and my husband for their support and encouragement throughout my study.

References

[1] Hong Liu et al," Skew detection for complex document images using robust borderlines in both text and non-text regions", Science Direct, Pattern Recognition Letters, 2009.
[2] Vladimir Y. Mariano et al, "Detection of text marks on moving vehicles", Proceedings of the Seventh international conference on Document Analysis and Recognition, IEEE 2003.
[3] Mehdi Felhi, Nicolas Bonnier, Salvatore Tabbone," A robust skew detection method based on maximum gradient difference and R-signature", 18th IEEE International Conference on Image Processing 2010.
[4] Li Sun et al, "A Novel Text Detection and Localization Method Based on Corner Response", IEEE International Conference on Machine Vision, 2009.
[5] Yang Cao, Heng Li," Skew Detection and Correction in Document Images Based on Straight-Line Fitting", IEEE-International Conference on Image processing,2009.
[6] Katherine L. Bouman," A Low Complexity Sign Detection And Text Localization Method For Mobile Applications", IEEE Transactions On Multimedia, Vol. 13, No. 5, October 2011.
[7] Andrej Ikica, Peter Peer, "An improved edge profile based method for text detection in images of natural scenes" , IEEE Transactions On Computer vision, October 2010.
[8] Kai Chen et al, "A Learning Based Text Detection Method in Camera Images", IEEE, International Conference on Computer Vision, 2011.
[9] Adam Coates, "Text Detection And Character Recognition in Scene Images Using Unsupervised Feature Learning" IEEE, International Conference on Document Analysis And Recognition, 2011.
[10] P Shivakumara∗, G Hemantha Kumar, D S Guru And P Nagabhushan," A novel technique for estimation of skew in binary text document images based on linear regression analysis", Sadhana Vol. 30, Part 1, February 2005, pp. 69– 85.

[11] Dhanya.D, Ramakrishnan.A.G And Peeta Basa Pati, "Script Identification In Printed Bilingual Documents," Sadhana, Vol. 27, Part-1, Pp. 73-82, 2002.
[12] Mallikarjun Hangarge and B.V.Dhandra, "Offline Handwritten Script Identification in Document Images", International Journal of Computer Applications, Volume 4 – No.6, July 2010.
[13] M. C. Padma, "Script Identification From Trilingual Documents Using Profile Based Features", International Journal of Computer Science and Applications, Vol. 7 No. 4, pp. 16 - 33, 2010.
[14] Xiaodan Jia and Xinnian Wang, "A Novel Algorithm for Character Segmentation of Degraded License Plate Based on Prior Knowledge", scientific research foundation of Liaoning Province of China, 2009.
[15] Palaiahnakote Shivakumara et al, "A New Gradient based Character Segmentation Method for Video Text Recognition", IEEE Computer Vision and Pattern Recognition, 2010.
[16] Huadong Xia and Dongchu Liao, "The Study of License Plate Character Segmentation Algorithm based on Vertical Projection", IEEE 2011.
[17] Youngwoo Yoon et al, "Blob Extraction based Character Segmentation Method for Automatic License Plate Recognition System", IEEE 2011.
[18] Youngwoo Yoon et al, "Blob Detection and Filtering for Character Segmentation of License Plates", IEEE 2012.
[19] J. A. Vlontzos and S. Y. Kung , "Hidden Markov Models for Character Recognition, " IEEE transactions on image processing, vol. 1 , no. 4, October 1992.
[20] C.N. Anagnostopoulos, I. Anagnostopoulos, V. Loumos, and E. Kayafas, "A License Plate Recognition Algorithm for Intelligent Transportation System applications," IEEE transactions on intelligent transportation systems, vol. 7, no. 3, September 2006.
[21] Takayuki Kurozumi, Yoshikazu Shinza, Yukiko Kenmochi and Kazunori Kotani, "Facial Individuality and Expression Analysis by Eigenspace method based on Class features or Multiple Discriminant Analysis," IEEE Proceedings of International Conference on Image Processing, vol. 1, pp: 648-652, 1999.
[22] Md. Anwar Hussain, "Eigenspace Based Accurate Iris Recognition System," 2010 Annual IEEE India Conference (INDICON), pp:1-3, 2010.
[23] Million Meshesha and C.V.Jawahar, "Optical character recognition of Amharic documents," African Journal of Information and Communication Technology, vol. 3, no. 2, 2007.
[24] M. Pietikäinen et al., Computer Vision Using Local Binary Patterns, Computational Imaging and Vision 40,Chapter 2,Springer-Verlag London Limited 2011
[25] Chi Ho Chan," Multi-scale Local Binary Pattern Histogram for Face Recognition", Centre for Vision, Speech and Signal Processing School of Electronics and Physical Sciences University of Surrey, September 2008.
[26] Yi-Feng Pan, Xinwen Hou, Cheng-Lin Liu ,"A Robust System to Detect and Localize Texts in Natural Scene Images", Eighth IAPR Workshop on Document Analysis Systems, IEEE 2008.
[27] Gang Zhou, Yuehu Liu, Quan Meng, Yuanlin Zhang ,"Detecting Multilingual text in Natural Scene", 1st International Symposium on Access Spaces (ISAS), IEEE-ISAS 2011.
[28] T.Senthil kumar et al, "An improved approach for Character Recognition in Vehicle Number plate using Eigenfeature Regularisation and Extraction Method", International Journal of Research and Reviews in Electrical and Computer Engineering (IJRRECE) Vol. 2, No. 2, June 2012.