Number Plate Recognition for Indian Vehicles

Number Plate Recognition for Indian Vehicles

M. M. Shidore[†], *S. P. Narote*[††]

[†]*Vishwakarma Institute of Technology, University of Pune, Pune, India*
[††]*Sinhgad College of Engineering, University of Pune, Pune, India*

Abstract: This paper presents Automatic Number Plate extraction, character segmentation and recognition for Indian vehicles. In India, number plate models are not followed strictly. Characters on plate are in different Indian languages, as well as in English. Due to variations in the representation of number plates, vehicle number plate extraction, character segmentation and recognition are crucial. We present the number plate extraction, character segmentation and recognition work, with english characters. Number plate extraction is done using Sobel filter, morphological operations and connected component analysis. Character segmentation is done by using connected component and vertical projection analysis. Character recognition is carried out using Support Vector machine (SVM). The segmentation accuracy is 80% and recognition rate is 79.84 %.

Key words:
Vehicle number plate, Number plate extraction, character segmentation, character recognition.

1. Introduction

Automatic Number Plate Recognition (ANPR) system is an important technique, used in Intelligent Transportation System. ANPR is an advanced machine vision technology used to identify vehicles by their number plates without direct human intervention. It is an important area of research due to its many applications. The development of Intelligent Transportation System provides the data of vehicle numbers which can be used in follow up, analyses and monitoring. ANPR is important in the area of traffic problems, highway toll collection, borders and custom security, premises where high security is needed, like Parliament, Legislative Assembly, and so on. The complexity of automatic number plate recognition work varies throughout the world. For the standard number plate, ANPR system is easier to read and recognize. In India this task becomes much difficult due to variation in plate model.

The ANPR work is generally framed into the steps: Number plate extraction, character segmentation and character recognition. From the entire input image, only the number plate is detected and processed further in the next step of character segmentation. In character segmentation phase each and every character is isolated and segmented. Based on the selection of prominent features of characters, each character is recognized, in the character recognition phase. Extraction of number plate is difficult task, essentially due to: Number plates generally occupy a small portion of whole image; difference in number plate formats, and influence of environmental factors. This step affects the accuracy of character segmentation and recognition work. Different techniques are developed for number plate extraction.

Hao Chen et al [1] proposed the method in two steps. First, several candidates based on texture information similar to number plate are extracted. In the second step, auto-correlation based binary image and projection algorithm are used to verify the true candidate plate. Gisu Heo [2] developed number plate detection algorithm using group of lines forming rectangle at the plate boundary. Followed by this step is the vertical edge density algorithm to find out the plate area. Ozbay et al [3] devised smearing algorithm to locate the number plate. In Korean license plate extraction, Mei Yu et al [4] proposed vertical edge detection followed by size, shape filter for edge area and edge matching technique based on plate model. Farhad Faradji et al [5] first used Sobel vertical edge detection on the image. Next, vertical projection analysis was used to locate plate area. Fake candidates were deleted using compact factor, which evaluated the densest vertical edge area declaring true number plate.

Every character on detected number plates is segmented in character segmentation step. Segmentation techniques based on projection analysis, Hough transform, region growing are proposed in the literature. Xinagjian He et al [6] used horizontal and vertical projection analysis for character segmentation.

Yuangang Zhang et al [7] implemented character segmentation using Hough Transform. In this, horizontal edges of the plate area were decided initially, using Hough Transform, which helped to segment the characters with the large rotation. Characters were segmented using vertical projection analysis based on the prior knowledge of the plate model. Feng Yang et al [8] developed region growing algorithm for character segmentation. Shen Zheng Wang et al [9] used connected component analysis for character segmentation.

Fundamental issue in character recognition is shape analysis. For the recognition of segmented characters, algorithms use Pattern/ Template matching, Computational Intelligence, statistical classifiers.

Manuscript received February 5, 2011
Manuscript revised February 20, 2011

ANPR system consists of three modules; 1) Extraction of number plate, 2) Character segmentation, and 3) Character recognition. In this paper ANPR wok for Indian car is presented. Images are taken out with different illumination conditions, different background and orientation. Histogram equalization, median filter are used which take care of lighting and contrast problem. Sobel vertical edge detection and morphology is employed to locate the number plate. Projection analysis is used to segment the characters present on the plate. For recognition work SVM is used.

The rest of the paper is organized as follows: Section 2 explains the proposed algorithm for number plate extraction; Section 3 explains the algorithm used for character segmentation; Section 4 describes the number plate recognition algorithm using SVM. Experimental results are demonstrated in Section 5. Section 6 concludes the paper.

2. Number Plate Extraction

Number plate extraction is the key step in ANPR system, which influences the accuracy of the system significantly. The goal of this phase, given an input image, is to produce a number of candidate regions, with high probability of containing number plate and validate for true number plate.

2.1 Image Acquisition and Preprocessing

In this system a high resolution digital camera is used to acquire an image. Images are taken in different background, illumination conditions, and at various distances from the camera to vehicle. Images are resized to (1024 X 768). All the processing steps are executed on gray scale image. Preprocessing is mainly used to enhance the processing speed, improve the contrast of the image, and to reduce the noise in the image. In order to reduce the problem of low quality and low contrast in car images, images are enhanced by using histogram equalization on gray scale image.

2.2 Vertical Edge detection

The characters on number plate region contain abundant edges as compared to background area. This feature is employed for locating the candidate plate area from the input image. Sobel vertical edge detection is used to find out the regions which have high pixel variance value [10]. To extract candidate number plate area from the entire image, threshold is used to select rows which are having particular white pixel density. Fig. 1 shows the result of effect of using, Sobel vertical edge detection and threshold.

(a) (b) (c)

Fig. 1: (a) Gray scale image (b) Sobel vertical edge detection (c) Effect of threshold

2.3 Candidate Plate Area Detection

Morphological operations aim to remove unrelated objects in the image. Dilation and erosion are used to extract candidate plate areas from the entire image. Sometimes background areas may also get declared as candidate plate. Hence to remove the fake candidates, plate validation is done using the aspect ratio of the plate and horizontal cuts [11] in the number plate.

2.4 True Number Plate Extraction

After the detection of candidate number plate area, Bounding Box analysis is used to extract plate area from the original image. From the Bounding Box analysis, respective row and column indices of plate area are found out. Once the indices of number plate are known, the number plate is extracted from original gray scale image. The result is as shown in Fig. 2.

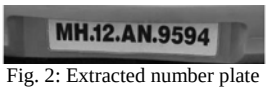

Fig. 2: Extracted number plate

3. Character Segmentation

Character isolation from the number plate region is the important step in ANPR system, which influences the accuracy of character recognition significantly. The goal of this phase, given the number plate image, is to segment all the characters, without loosing features of the characters. This phase consists of the sequence of operations as, character region enhancement, connected component analysis and projection analysis.

3.1 Character Region Enhancement

To enhance the character region, if histogram equalization or gray level scaling is used, noise may also be enhanced. For the correct character segmentation, only the character pixels need to be enhanced and the background pixels need to be weakened. For this ‘gray threshold’ [10] is implemented. Fig. 3 shows thresholding effect.

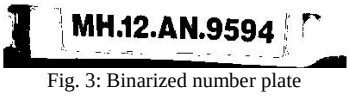

Fig. 3: Binarized number plate

3.2 Connected Component Analysis

To remove the noise, other than characters on the plate, connected component analysis is used. Each labeled matrix of 8-connectivity pixels is evaluated based on the area threshold. The result of connected component analysis and noise removal is as shown in Fig. 4.

(a) (b)

Fig. 4: (a) Connected component analysis (b) Noise removed

3.3 Vertical Projection Analysis

Character segmentation is implemented to isolate each character on the number plate. Vertical projection analysis is used to find the gaps between the characters. The characters are segmented based on the number of valleys in projection. Row and column indices of each character are recorded, and it is extracted from the original gray scale number plate. The vertical projection and segmented characters are shown in Fig. 5. This proposed work isolates every letter and digit on the number plate oriented horizontally in one row, along the width of number plate.

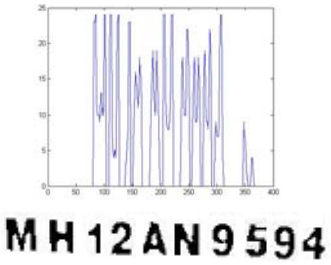

Fig. 5: Vertical projection analysis and respective segmented characters

4. Character Recognition

The character recognition phase consists of two steps: 1) Character normalization and feature extraction, 2) Character classification using Support Vector Machine [12].

4.1 Character Normalization

Segmented characters have very much variation in size. In this phase, all the characters are normalized to predefined height (Vertical Length) in pixel. As the characters always have variable width (Horizontal Length), each character image is normalized to a size of 32 X 32, by image mapping technique.

4.2 Feature Extraction

The goal of feature vector is to define distinguishing features of the characters. Selecting the most relevant feature of each character can not only facilitate data visualization and data understanding, but also reduce the measurement, storage requirements, training and utilization time, particularly when the features are redundant.

Initially, the centroid of the character image is determined. With respect to centroid, number of transitions along the axes, 0 to 1 and 1 to 0, up to the boundary of character are counted. Transitions are specified for axes with predetermined angles. Fig. 6 shows the nature of axes towards boundary, counting number of transitions. By observing strokes of each character, 13 different angles are decided to count the transitions.

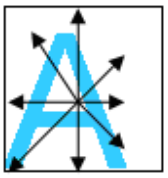

Fig. 6: Input character image overlaid with nature of axes

4.3 Number Recognition

The character recognition work is carried out using SVM as classifier. The car number plate consists of combination of 0 to 9 digits and alphabets of A to Z. All the digits as well as alphabets are assigned a class label. For multi-class classification, either one-against-one approach is used or one-against-all is preferred. In the proposed work recognition is achieved using one-against-all approach [13]. The polynomial kernel is used for classification. The database consists of 250 images. 60 % of the data is used for training and remaining 40% is used for testing.

5. Experimental Results and Performance Analysis

A database consists of different sized JPEG colored images. Images are resized to 1024 X 768. Total 250 images are used to test the algorithm. The images are taken with different background as well as illumination conditions. Experiments show that the algorithm has good performance on number plate extraction, and character segmentation work. It can deal the images correctly, with noise, illumination variance, and rotation to $\pm 5^0$. This work is implemented using MATLAB.

Table 1 illustrates number plate extraction and character segmentation success rate.

Table1 Performance of number plate extraction and character segmentation.

Algorithm	Total No. of Images	Success Rate (%)
Number Plate Extraction	250	85
Character Segmentation	214	80

Deep shadows and reflections have an impact on number plate extraction work. Because of uneven illumination, stained number plates, true number plates could not get

correctly extracted. Failure in character segmentation was mainly because of merging of characters on number plate, stained number plates, orientation of the image and poor illumination.

Character recognition work is done on 10 digits (0 to 9) and 26 alphabets (A to Z). The recognition rate achieved using 'polynomial' kernel is 79.84 %. The false recognition is due to similarity in the character shape, e.g. 6 and B, 5 and S etc. Recognition can be increased if two groups of SVM's are designed separately, one for digits and other for alphabets.

6. Conclusions

An algorithm for vehicle number plate extraction, character segmentation and recognition is presented. Database of the image consists of images with different size, background, illumination, camera angle, distance etc.

The experimental results show that, number plates are extracted faithfully based on vertical edge detection and connected component algorithm, with the success rate of 85%. Character segmentation phase using connected component analysis and vertical projection analysis works well with the success rate of 80%. The success rate achieved for character recognition is 79.84%.

References

[1] Hao Chen, Jisheng Ren, Huachun Tan, Jianqun Wang, " A novel method for license plate localization", 4th Proc. of ICIG 2007, pp. 604-609.

[2] Gisu Heo, Minwoo Kim, Insook Jung, Duk Ryong Lee, Il Seok Oh, "Extraction of car license plate regions using line grouping and edge density methods", International Symposium on Information Technology Convergence, 2007, pp. 37-42.

[3] Serkan Ozbay, Ergun Ercelebi, "Automatic vehicle identification by plate recognition", Proc. of PWASET, vol. 9, no. 4, 2005, pp. 222-225.

[4] Mei Yu and Yong Deak Kim, "An approach to Korean license plate recognition based on vertical edge matching", IEEE International Conference on System, Man and Cybernetics, 2000, vol.4, pp. 2975-2980.

[5] Farhad Faradji, Amir Hossein Rezaie, Majid Ziaratban, "A Morphological based License Plate Location", ICIP, 2007, pp. I 57-I 60.

[6] Xiangjian He et al, "Segmentation of characters on car license plates", 10th Workshop on Multimedia Signal Processing, Oct. 2008, pp. 399-402.

[7] Yungang Zhang, Changshui Zhang, "A New algorithm for character segmentation of license plate", Proc. of IEEE Intelligent Vehicles Symposium, 2003, pp. 106-109.

[8] Feng Yang, Zheng Ma, Mei Xie, "A Novel approach for license plate character segmentation", ICIEA, 2006, pp.1-6.

[9] Shen-Zheng Wang, His-Jian Lee, "Detection and recognition of license plate characters with different appearances," Proc. of 16th International Conference on Pattern Recognition, vol.3, 2003, pp. 979-983.

[10] Rafel C. Gonzalez, Richard E. Woods, "Digital Image Processing," *2nd edition, Prentice Hall, Inc.*, 2002.

[11] Tran Dur Duan, Duong Anh Duc, Tran Le, Hong Du, "Combining Hough transform and contour algorithm for detecting vehicle license plate", Proc. of International Symposium Intelligent Multimedia, Video and Speech Processing , 2004, pp. 747-750.

[12] Christopher J. C. Burges, "A tutorial on Support Vector Machine for pattern recognition," Bell Laboratories, Lucent Technologies, 1998.

[13] S. R. Gunn, "Support Vector Machines for classification and regression, technical report," University of Southampton, 1998.

Mrunal M. Shidore was born in India in 1964. She received the B.E. (Electronics) degree from University of Pune, Pune in 2003. She is currently working as a Lecturer in Electronics Engineering, with the Vishwakarma Institute of Technology, Pune. Her research and teaching interests include image processing.

Sandipan P Narote was born in India in 1975. He received the B.E.(E&TC), M.E.(Elect.) from Dr. Babasaheb Ambedkar University Aurangabad and Ph.D.(E & TC) from Swami Ramanand Tirth Marathwada University Nanded in 2000, 2002 and 2010, respectively. He is currently a Professor of Electronics and Telecommunication Engineering with the Sinhgad College of Engineering, Pune, India. His research and teaching interests include signal processing, image processing and biometrics.