

THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

A Survey on OCR Based Number Plate Recognition for Indian Conditions

P. B. N. Chakravarthy

M.Tech Student, Computer Networks Information Security & JNTU
School of Information Technology, India

E. Jagadeeshwar Rao

Assistant Professor, Software Engineering, JNTUH, School of Information Technology, India

Abstract:

OCR based number plate recognition system automatically recognizes the license number of vehicles. This system has algorithms like: 'Feature based number plate Localization' for locating the number plate, 'horizontal and vertical projection' for character segmentation Template method to normalize the characters and OCR for character recognition. The other main feature of this system is connecting to the database. This survey presents different types of techniques to recognize characters and different techniques to improve accuracy.

Keywords: OCR, Image Processing, Feature Extraction, Number recognition

1. Introduction

OCR based number plate recognition system can able to upload and select image of vehicles and recognizes the license number on that image. This system usually performs pre-processing over the images obtained from the camera placed over the traffic. These systems use the Optical Character Recognition (OCR) technique to read numbers on number plate (image).

There are so many Number plate Localization algorithms but several different procedures involved in one algorithm hence they require long computational time. The results obtained after applying algorithm are dependent on the image quality. If it is a noisy picture then the result is not accurate. So, along with the algorithm clear picture is also important. If the image is blurring then it is impossible to recognize the number correctly as on number plate. Hence we should careful while taking a picture but in the case of traffic the vehicles are moving so we have to use motion picture capable cameras. In foreign countries detection of the number plate is done on the basis of attributes such as size of plate, border, colour of border, colour of characters and font on number plate.

2. Literature Survey

In the Existing System Optical Character Recognition is the not primary technique to analyze the image. It uses merging of several algorithms with more computation time. Existing Automatic vehicle Number Plate Recognition was invented in the year 1976 at the Police Scientific Development Branch in UK for safety and security.

2.1. Various Methods

There are different approaches in number plate detection some of them are:

2.1.1. Hierarchical approach and overlapped portioning technique

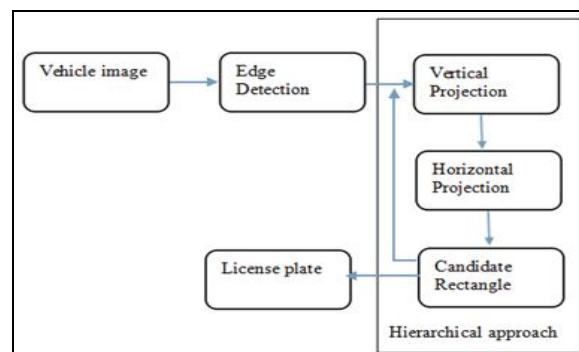


Figure 1: hierarchical approach and overlapped portioning technique

2.1.2. Using edge/boundary detection method

The Edge/Boundary-based extraction method uses Hough transform (HT). The Hough transform has the advantage of detecting straight lines and lines with up to 30° inclination. Hence we can detect numbers even the image is inclined to some angle (<30°).

2.1.3. Using color and character features

Based on the hue, lightness and saturation the pixels in the input image are classified into 13 categories. Among the 13 categories we select one or more categories for detection of number plate on image.

2.1.4. Connected component labeling

Connected components labeling method used to scan an image and group its pixels into components. Each component consists of set of pixels that have similar pixel intensity values, components are also connected to each other component based on one similarity. If all the groups have been determined then assign gray or a color (color labeling) for each pixel of each component. It works on black & white or gray or binary images.

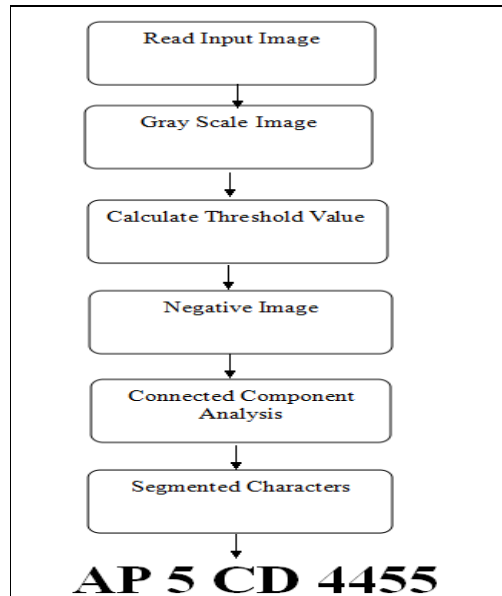


Figure 2: process of connected component labeling

2.2. Proposed Project Work

From the survey that was conducted on different existing systems techniques the new project “OCR Based number plate recognition for Indian Conditions” consists of following steps:

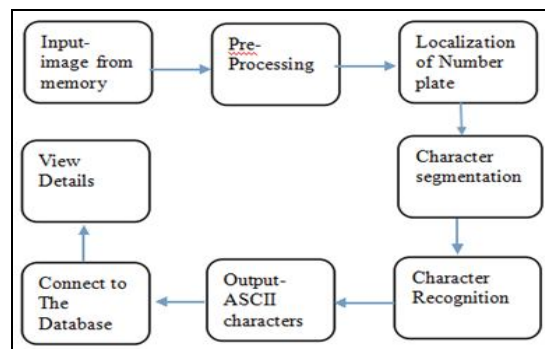


Figure 3: steps in proposed system

2.2.1. Input Image

The End user from the user module has to upload images from the memory (either from disk or from computer memory) and should select any one of them to process further. In India there are two types of license plates are used. One is for commercial vehicles and other is for Private Vehicles. The license plate of commercial vehicle has a yellow background for the plate with black scripts on it. The license plate for private vehicles has a white background with black script on it. In India the first two letters on license plate indicate state code and it is followed by two-digit numeric code of the district in that state.

2.2.2. Pre-Processing

Pre-Processing is the first step in the process of recognizing numbers on number plate. To recognize numbers on image of vehicle, convert the image into either gray scale or to binary image. In this pre-processing we use some techniques to improve the image quality and reduce the noise of the image. The Proposed project uses two basic techniques. They are: Edge Detection and Binarization.

- *Edge Detection:* Edge Detection technique is used to identify points where brightness changes sharply and continuously in the image. Edges are boundaries between two different regions. Some of the edge detection algorithms are: Robert's edge detector method, Prewitt edge detector method, Frie Chen edge detector method, Canny edge detector and Sobel edge detector method. Among them all Canny Edge Detector is used in proposed system since it doesn't involve more than three steps and it results only high brightened areas of the image. For Indian Conditions it is better to use Prewitt Edge Detector Method since it results in less and more brightened areas.
- *Image Filtering:* Image filtering method is used to remove noise. But it can also increase brightness and contrast of image along with removing noise. For example, the values of R,G,B for a pixel are (R,G,B)=(30,127,254) then R',G',B' values are (R+G+B)/3 which is 137. A Gray scale conversion is obtained from each pixel (p, q) of 24-bit color value with R, G and B components are separated and the Gray value of that pixel is calculated using the formula: $\text{Gray}(p, q) = 0.07 * R(p, q) + 0.05 * G(p, q) + 0.5 * B(p, q)$ of 8-bit value, which means we are converting image to inverted image of more black colour.



Figure 4: Originally uploaded image

Figure 5: Grayscale image

Figure 6: filtered Image

2.2.3. Number Plate Localization

Number Plate localization locates number plate position from the output of the previous step using appropriate algorithm. This step recognizes the location of number plate in captured image so that it will be easier to recognize numbers and characters by using only that part of the captured image. It gives more sharp and clear edges in the image by applying algorithms.

2.2.3.1. Number plate Detection:

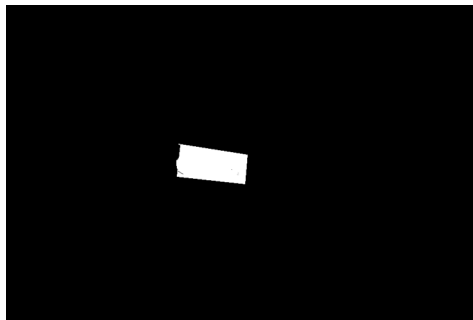


Figure 7: Number plate location

Localization is difficult if the images are of different colours, size. Various techniques were introduced recently for efficient detection of number plate regions from images. If the image is captured outside then there are variations in conditions like lighting, speed of wind, pollution and motion, which makes localization of true license plate regions as highly difficult. This difficulty can be reduced using Window filtering method, based on window filtering method, the size of the window is fixed to particular dimensions and it is eventually scans the total image to recognize particular place where the window consists of number plate only.

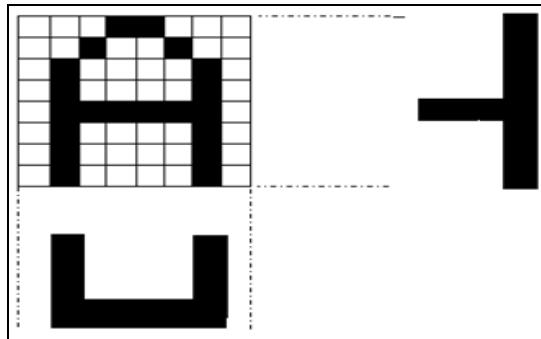
2.2.3.2. Extracted Number Plate



Figure 8: Extracted Number Plate from image

2.2.4. Character Segmentation

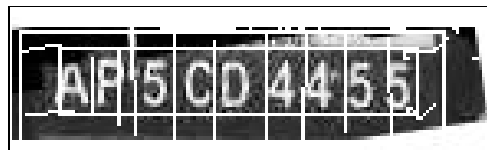
Based on horizontal and vertical projections on the image the black regions or white regions on the image are taken as a part of the image. Horizontal projection results in a group of black pixels ranges from minimum to maximum. Vertical projection also ranges from minimum to maximum. By combining the two resulted graphs we can able to draw a template matrix where the only black pixels are to be taken as one or white pixels to be one. The character clipper is then used on the template matrix to recognize character. Then by using feature extractor each character is segmented where one character to one template matrix.



Horizontal projection is the total number of all black-pixels along a column while vertical projection is the total number of all black-pixels along a row of a character image. Hence horizontal projection of a character results columns and vertical projection results rows. For example, if there is v number of rows and va is value of v , then vertical projection:

$V = \text{sum of } (v[i]*va)$, where I ranges from 1 to v .

2.3. Image after segmentation can be as follows



The Segmented characters all of same size, width and height.

2.3.1. Character recognition

To recognize Characters we use Optical Character Recognition method. In OCR numbers, characters on image are converted into editable text. OCR can extract characters from image using two techniques. They are: Template matching and Artificial Neural Networks.

The OCR recognition algorithm depends upon a set of segmented characters and properties of those characters. It compares the characters in the segmented image files with the characters in the pre-defined set. Generating the pre-defined set is simple. Pre-defined set consists of an image file with the specified characters with specific font and a text file containing characters in that image file is created.

Once the pre-defined set has been read from the image file and its properties are recognized, it can be written into a "context.css" file. This file stores the properties of the pre-defined characters and it is accessible throughout the project. It reduces accessing time of pre-defined set for each segmented character using java in the program.

After character segmentation the each character is resized to 9X12 or 12X18 or 24X32 etc. based on the characters we have in our memory to compare. Character segmentation is very important in order to perform character recognition with good amount of accuracy (correctly). Sometimes character recognition is not accurate due to error in character segmentation.

- **Template Matching:** It compares the output images of previous step with the Characters that are already in the memory as files. Here OCR extracts the connected components from each line by scanning from the upper right corner of each line. It removes touching intervals of black pixels or white pixels from the segmented image until nothing more connected can be found. In the extraction step, it further looks entire segment to see if there are possible "extra parts", such as the dot on

'j'. Extra parts are hanging directly above or below the component. It ignores components that are above or below the extracted component, as when the tail of a 'y' extends below the letter preceding it.

2.3.2. Connect to database

At last the characters that recognized are transferred to the database to view results on owner details. The Database should be connected to the RTO's database but it results security problems so to reduce those security problems we are connecting to the internal database.

3. Conclusion and Future Scope

In this paper we discussed number plate detection problems in different algorithms. This system is useful in reducing crime rate if it is introduced at traffic and at high crime rate areas. Though plenty of algorithms were introduced, still user is facing problems with the recognition of characters. Finally I conclude that there is a necessity to improve OCR algorithm and this system should be introduced into the government of transportation or to RTO so that we can get correct details of the vehicle owner.

4. Acknowledgment

This work is supported and made under the esteemed guidance of Mr. E Jagadeeshwar Rao, who is a lecturer and head of the department of Software Engineering at School of IT, Jawaharlal Nehru Technological University, and Hyderabad.

5. References

1. Automatic Number Plate Recognition (ANPR) System for Indian conditions, Prathamesh Kulkarni (Student Member, IEEE), Ashish Khatri, Prateek Banga, Kushal Shah, University of Pune, Dept. of Electronics and Telecommunication, India. 2009
2. Srivastava R: Transformations and distortions tolerant recognition of numerals using neural networks, ACM Annual Computer Science Conference, San Antonio, Texas, USA, 1991.
3. Zhang Y., Zhang C.: New Algorithm for Character Segmentation of License Plate, Intelligent Vehicles Symposium, IEEE, 2003.
4. Yo-Ping Huang, Shi-Yong Lai, Wei-Po Chuang, "A Template-Based Model for License Plate Recognition", IEEE International Conference on Networking, Sensing & Control, March 21-23, 2004.
5. Indian Central Motor Vehicle Act, Bear Act, Rule no.49, 50