International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-1, Issue-3, July 2011

Morphology Based Approach To Recognize Number Plates in India

Phalgun Pandya, Mandeep Singh

Abstract—**Automatic number plate recognition is a real-time embedded system which automatically recognizes the license number of vehicles. Such systems require the localization of number plate area in order to identify the characters present on it. This paper presents an approach based on simple but efficient morphological opening and closing operations for localization of Indian number plates. After localization the skew correction of the number plate is also done for effective segmentation of characters. The number plate skew correction greatly affects the accuracy of the character extraction. The character extraction is based on template matching approach. Our proposed algorithm has been tested on 100 samples and is found to be robust to detect vehicle license plates as well as extracting the numbers with an overall accuracy of 90%.**

Keywords— **Character Recognition, Template Matching, Morphology, LPR, ANPR.**

I. INTRODUCTION

Every country uses specific vehicle identification system for the control of traffic, traffic surveillance, monitoring against illegal activities, security control of restricted areas, traffic law enforcements, toll collection and parking management etc and as such India also has its own system of assigning unique numbers to vehicles. These unique numbers plates are assigned to the vehicles by RTO (Regional Transport Office).These plates in general are easily readable by the human beings because of very high level of intelligence but when it comes to do the same using machines, many effects such as illumination, blur, background and foreground color etc. pose a problem. Also the License plate recognition (LPR) in India is difficult in the sense that the traffic rules are hardly followed. Since number plate standards are not strictly practiced in India, a large amount of variations are obtained in parameters like, size of number plate and characters, location of number plate, type of font used (standard font is Arial Black), background (white for non commercial vehicles and yellow for commercial vehicles) and foreground color (black for commercial and non commercial vehicles), etc. which makes the task of number plate localization all the more difficult. Main aim of this research paper is to implement a method efficient in recognizing license plates in Indian conditions. Our work is not restricted to car but is expanded to many types of vehicles like motor cycle (in which size of license plate is small), transport vehicles which carry extra stylized characters and soiled license plate. Our proposed algorithm is found to be robust to detect vehicle license plates as well as extracting the numbers from them. In India we have various kinds of number plates. Old number plates following 1939 series as well as the vehicles following new number system. The new format license plates can be of lengths 8,9,10. A typical example of an Indian license plate (for car) is shown in the fig. 1. with the significance of each character. Another variant of an Indian license plate which is mostly used for two wheelers is also shown below it.

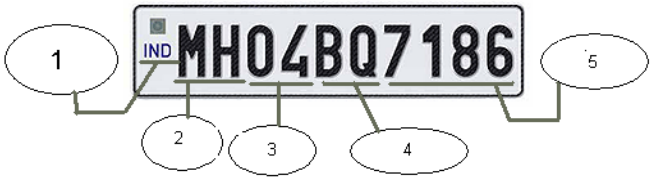


Figure 1 Format of Indian License plate (Image courtesy of Google Images)

The latest license plate format is shown above.

1. Country Code
2. State Code
3. District Code
4. Type of Vehicle (car, two wheeler, commercial etc.)
5. Actual Registration Number

Figure 2 Another Valid License Plate format (image courtesy of Google Images)

There are a number of algorithms proposed for number plate localization such as multiple interlacing and transform domain filtering. In case of multiple interlacing algorithm [1], horizontal edge detection and vertical edge detection is performed separately on input vehicle image. Then resulting images are then added to get an image which generates co-ordinates of number plate. This approach cannot be employed for Indian number plates because they do not necessarily have a border which is found to be essential condition for this algorithm. Transform domain filtering [1] is another approach in which high frequency area of input image is taken out as number plate. This algorithm is also not suitable for Indian conditions because of the presence of other

Manuscript received June 21, 2011.
Phalgun Pandya, Department of Electrical & Instrumentation Engineering, Thapar University, Patiala, 147001, Punjab, INDIA.
Mobile No. +919501313917, (e-mail: phalgun.jecrc2008@gmail.com).
Mandeep Singh, Asst. Prof., Department of Electrical & Instrumentation Engineering, Thapar University, Patiala, 147001, Punjab, INDIA.
Mobile No. +919815605616, (e-mail: mdsingh@thapar.edu).

characters or character like regions in input image. Velappa Ganapathy [2] proposed license plate localization and recognition system for vehicles in Malaysia based on a combination of morphological processes with a modified Hough Transform approach for skew correction. Since Hough Transform is computation intensive, it eats up vital time required for further computation. Some algorithms [3]-[4] based on morphology have also been proposed but they are restricted to localization only.

Our proposed method applies basic mathematical morphology operations like 'opening' and 'closing' along with very simple heuristics for the localization of the license plates and character recognition technology to extract the numbers from the localized license plate.

II. Related Definitions And Mathematics

Here are some important definitions used throughout the paper.

1) ANPR- Automatic Number Plate Recognition

2) LPR- License Plate Recognition

3) Mathematical morphology- It is a topological and geometrical based approach for image analysis which provides powerful tools for extracting geometrical structures and representing shapes in many applications.

The following operations form the basis of mathematical morphology-

3.a) Dilation-

$$(F \oplus B)(x, y) = \max\{F(x - s, y - t) + B(s, t)\} \quad (1)$$

3.b) Erosion -

$$(F \ominus B)(x, y) = \min\{F(x + s, y + t) - B(s, t)\} \quad (2)$$

3.c) Opening

$$F \circ B = (F \ominus B) \oplus B \quad (3)$$

3.d) Closing

$$F \bullet B = F \oplus B) \ominus B \quad (4)$$

In all the above operations F(x,y) is a grayscale image and B(x,y) is the structuring element.

4) Structuring Elements- In mathematical morphology these are represented as matrices. Structuring element is a characteristic of certain structure and features to measure the shape of an image and is used to carry out other image processing operations. The shape and size of the structuring element (SE) plays crucial role in image processing and is therefore chosen according to the condition of the image and demand of processing.

Typical structuring elements are shown below:-

1	1	1	1	1	1
1	1	1	1	1	1

A 2*5 'Rectangular' Structuring Element

			1			
	1	1	1	1	1	
	1	1	1	1	1	
1	1	1	1	1	1	1
	1	1	1	1	1	
	1	1	1	1	1	
			1			

A 'disk' shape structuring element having radius 3

III. Proposed Method For Recognization Of License Plates

Generally the text written on the number plates has high contrast as compared to the background and based on this property of text a localization technique has been proposed.

The work is divided into four parts:-

- Binarization
- Filtering Candidate Regions
- Verification of the Localizes candidate
- Number extraction and verification

This work aims on gray intensity based license plate extraction and hence begins with color to gray conversion.

A. *Binarization*

Step 1. Let the input grayscale image be named as F(x,y).

Step 2. The image was then filtered using median filter to create a blurred image (to prevent over segmentation)[5] and the difference image[5,6] was obtained by subtraction of opened grayscale image (Fo(x,y)) from a closed grayscale image(Fc(x,y)) using equations 3 and 4.

Step 3. The Difference image is then closed and opened using 5*5 size structuring element.

Step 4. The resulting image is then binarized using global threshold value [1]-[5]. This completes the Binarization process.

B. *Candidate Region Filtering*

Since a number plate can be assumed to have larger lengths as compared to their width and also they are mostly horizontally aligned as compared to other regions a region based filtration technique can be applied to remove non-candidate regions.

Step 4. Label the binarized image.

Step 5. Filter the regions such that
Area(R)> Total Image area/400 [6]-[7]
Aspect Ratio(R) > 2, and
den(R)>0.1
where Aspect Ratio=width/height
and den(R)=Area of R/(w(R) * h(R))
R defines the region of interest; w(R) and h(R) are the bounding box width and height respectively;

It is possible that the number plate regions thus extracted can be fragmented in parts and hence these parts should be merged in order to correctly localize the plate. The basis for merging fragments is that if the number plate is recognized but in fragments then both the fragments should have similar heights and orientations just like fragmented text lines [7].

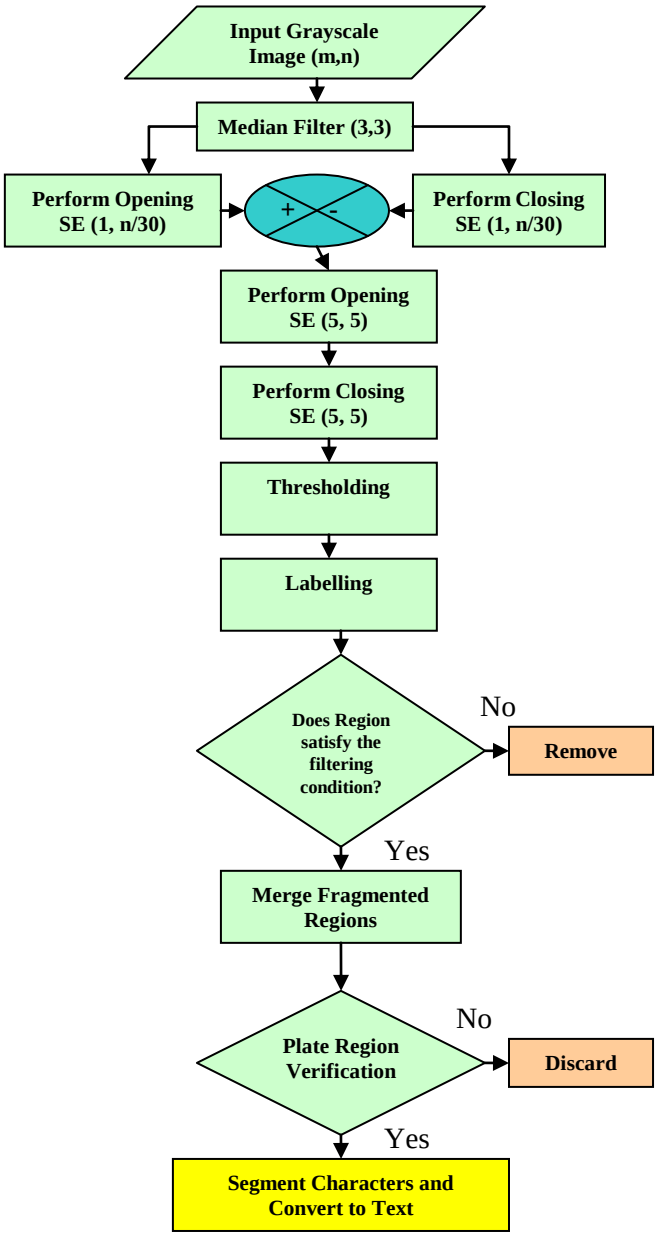
Step 6. Merge the i^{th} and j^{th} regions based on the following equations

$$d\theta(R_i, R_j) = min(|\theta_{Ri} - \theta_{Rj}|, |\theta_{Ri} - \theta_{Rj} + 360^\circ|, |\theta_{Ri} - \theta_{Rj} - 360^\circ|) < 10^\circ \quad (5)$$

where θ_{Ri} and θ_{Rj} are the major orientations of R_i and R_j, respectively. The threshold 10° is not related to the sizes of image and font.

Let Cen_{Ri} and Cen_{Rj} be the centroids of R_i and R_j respectively. The third criterion requires Cen_{Ri} and Cen_{Rj} being closer to each other; that is,

$$|Cen_{Ri} - Cen_{Rj}| < 2(h_{Ri}\theta_{Ri} + h_{Rj}\theta_{Rj}) \quad (6)$$



Flowchart of the Proposed Algorithm

Let L_R be the longest axis of R denoted by this equation:

$$y = m_R x + b_R$$

Then, the distance between the major axes L_{Ri} and L_{Rj} of R_i and R_j can be defined as follows

$$dL(Ri, Rj) = \{ |yCen_{Rj} - m_{Ri}\, xCen_{Rj} - bRi| / 2*sqrt(1 + m^2_{Ri}) \} - \{ |yCenRi - mRj\, xCen_{Ri} - b_{Rj}| / 2*sqrt(1 + m^2_{Rj}) \} \quad (7)$$

Such that dL<5;

where x_{CenR} and y_{CenR} are the coordinates of Cen_R in the x and y directions respectively.

If the regions satisfy the equations 5, 6 and 7, the respective i^{th} and j^{th} regions should be merged.

Since the font size in a number plate is quite large, the value of dL may have to be increased.

C. *Verification of localized candidate*

Even after this, if false number plate location is present, it is discarded by using character count. The number of characters in finalized number plate areas is calculated. If number of characters is less than four, then it is discarded. After applying these steps on an input image, all other unwanted data except number plate is removed. Number plate co-ordinates are applied on an input gray-scale image and number plate is extracted. Extracted number plate is binarized using Otsu's method [8] to enhance its quality.

The fig. 4 to 9 shows the stages of the above mentioned steps.

Figure 3 Grayscale Image

D. *Number extraction and verification*

After the number plate has been extracted, the extracted sub-image is then passed through character recognition module.

Step 7. Since the region shown in fig. 8 is inclined at some angle, it needs to be corrected for increasing the accuracy of character extraction. This was done by rotating the extracted binary image by an angle equal to orientation angle of the region in fig. 8 but in opposite direction. The corrected image is shown in fig. 11

Figure 4 Grayscale Image after morphological operations

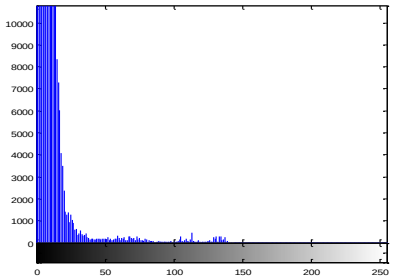


Figure 5 Histogram of figure 4

Figure 6 Binarized Image

Figure 7 Labelled image of figure 6

Figure 8 Localized Number Plate Region

Figure 9 Number Plate region in grayscale image

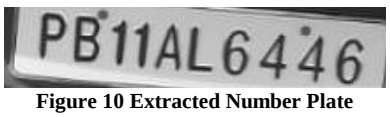


Figure 10 Extracted Number Plate

The difference between the fig.6 and fig.8 clearly shows that the unwanted regions are eliminated by the constraints applied and only the license plate region clears the filtration process. The license plate region has been highlighted in the original grayscale image by a 'cyan' colored rectangular box in fig. 9. and the extracted plate has been shown in fig. 11.

Step 8. Remove unwanted areas such as the regions belonging to plate fixing bolts and boundary areas. Segment each character from the binarized image [9].

Figure 11 Binarized number plate

Figure 12 Cleaned Binary Image

Step. 9 The segmented characters were then passed through template matching process [10]-[11]-[12]. Each segment is matched with a pre-loaded collection of templates (Arial Black and Times New Roman fonts have been used) and the segment having highest correlation score is considered to be the best match.

Figure 13 Segmented Characters from the plate

The process continues until all the segments have been identified.

Step 10. Since the accuracy of template matching is highly dependent on the contrast and illumination conditions, some similar characters can be misinterpreted. The characters such as 'B' '0' '1' 'G' can be misinterpreted as '8' 'O' 'I' 'C or 6' having similar characters which leads to lowering of Recognition efficiency[14].

Step 11. The number thus extracted can stored in an array and can be verified on the basis of rules [13]-[15] followed in India (fig. 1).

In the case of Indian number plates, the length can be 8, 9 or 10. Rules corresponding to the number plate lengths are as follows

- If the number of segmented characters counts to 7, the number plate follows 1939 series having first three characters as alphabets and rest as numbers.
- If the number of character counts to 8, the first two characters should be alphabets and rest should be numbers.
- If the number of character counts to 9, the first, second and fifth characters should be alphabets and rest should be numbers.
- If the number of segments counts to 10, the first, second, fifth and sixth characters will be alphabets and rest will be numbers.

- Exceptions- The foreign vehicle number plates contain third and fourth characters as alphabets and rest as numbers. Also for the case of Delhi not following 1939 series the fourth character is an alphabet containing information about the type of vehicle.

Step 12. At last when all the characters are verified, the number is stored to a text file.

IV. Experiments And Results

The tests were made on 100 images taken with the help of a digital camera, 98% of the number plates were localized correctly and 2% images resulted in the localization of number plates along with unwanted non candidate regions. The correctly localized 98 plate images when passed through Character Recognition module provided 91% accurate results. The remaining 9% were either fully or partially inaccurate.

The following images show the result of the proposed method.

Figure 14 Result of figure 13 = PB 11 AL 6446 (Printed in text file)

Figure 15 Localization of unwanted region along with License plate

Figure 16 Result of figure 15 = PB 10 CM 7057 (Printed in text file)

Figure 17 Extracted License Plate

Figure 18 Result of figure 17 = PB 11 AM 6811 (Printed in text file)

Figure 19 Extracted License Plate

Figure 20 Result of figure 19 = DL 7C 5902 (Printed in text file)

Figure 21 Extracted License Plate

Figure 22 Result of figure 21 = CH 03 B 5498 (Printed in text file)

Figure 23 Extracted License Plate

Figure 24 Result of figure 23 = CH 04 C 6009 (Printed in text file)

The following resulted in correct localization but false number recognition.

Figure 25 License plate with wrong result CH 04 O 2510 (error due to variation in font types)

Figure 26 Extracted License Plate

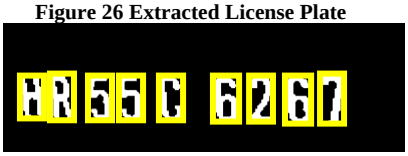


Figure 27 Result of figure 26 = BR-55-C-5267

Recognization Accuracy (%) =100*(1-E/T)

E=total error in character reconization; T=total number of characters in the number plate;

The results of fig. 25 and fig. 26 show that the stylized as well as inclined characters create ambiguity during template matching and hence pose a problem for correctly recognizing the vehicle license plates.

TABLE I
ASSESSMENT OF RESULTS

Results						
Number plate image	**Localization status**	**Segmented characters**	**Character recognization results**	**Character Recognization accuracy**		
				Error in alphabet	**Error in number**	**Accuracy (1-E/T)*100**
PB11AL6446	SUCCESS	PB11AL6446	PB11-AL-6446	NIL	NIL	100%
PB10CM7057	SUCCESS	PB10CM7057	PB10-CM-7057	NIL	NIL	100%
PB11AM6811	SUCCESS	PB11AM6811	PB11-AM-6811	NIL	NIL	100%
DL7C5902	SUCCESS	DL7C5902	DL7C-5902	NIL	NIL	100%
CH03B5498	SUCCESS	CH03B5498	CHO3-B-5498	NIL	1	((8/9)*100 = 88%
CH04C6009	SUCCESS	CH04C6009	CHO4-C-6009	NIL	1	((8/9)*100 = 88%
HR03E9707	SUCCESS	HR03E9707	HR03-E-9707	NIL	NIL	100%
PB11AK3244	SUCCESS	PB11AK324	PB11-AK-324	NIL	1	((9/10)*100 = 90%
PB11AS3761	SUCCESS	PB11AS376	P811-A5-376	2	1	((7/10)*100 = 70%
HR09L2071	SUCCESS	HR09L2071	HR09-L-2071	NIL	NIL	100%
HR55C6267	SUCCESS	HR55C6267	BR55-C-5267	1	1	((7/9)*100 = 77%
CH04A2519	SUCCESS	CH04A2519	CH04-O-2519	1	1	((7/9)*100 = 77%

V. CONCLUSIONS

A morphology based approach for localization of Indian license plates along with the number recognition based on template matching has been proposed. In this approach, number plate located at any corner of image can be recognized. Number plates having variations in background as well as font can be easily localized. Unwanted conditions such as screws and unwanted text on number plate which create problem for localization are treated suitably and taken into consideration. This algorithm has been found to be 90% accurate overall. The major sources of error were due to misinterpretation of characters in the number plate caused by extreme variation in the dimensions, overlapping and style of the characters and low contrast in images. However these errors can be aptly rectified by using more adaptive thresholding methods and advanced techniques such as trained Neural Networks.

REFERENCES

[1] Prathamesh Kulkarni, Ashish Khatri, Prateek Banga, Kushal Shah "A Feature Based Approach for Localization of Indian Number Plates" *IEEE International Conference on Information Technology, pp: 157-162*, 2009.

[2] Velappa Ganpathy and Wen Lik Dennis LUI "A Malaysian Vehicle License Plate Localization and Recognition System" *Journal of Systemics, Cybernetics and Informatics, ISSN 1690-4524, Vol. 6, No. 1*, 2008.

[3] P. V. Suryanarayana, Suman K. Mitra, Asim Banerjee and Anil K. Roy "A Morphology Based Approach for car license Plate" *IEEE Indicon 2005 Conference, Chennai, India, 11-13 Dec.* 2005.

[4] C.Nelson Kennedy Babu, Krishnan Nallaperumal "A License Plate Localization using Morphology and Recognition" *In Proceedings of India Conference, Volume 1, pp.34 – 39*, 2008.

[5] Yassin M. Y. Hasan and Lina J. Karam, "Morphological text extraction from images" *IEEE Trans. Image Process. 9(11), pp. 1978–83*, 2000.

[6] Jui-Chen Wu · Jun-Wei Hsieh · Yung-Sheng Chen. "Morphology Based text line Extraction" *Machine Vision and Applications , pp.195–207*, 2008.

[7] Kaushik Deb, Andrey Vavilin, Jung-Won Kim, and Kang-Hyun Jo "*Vehicle License Plate Tilt Correction Based on the Straight Line Fitting Method and Minimizing Variance of Coordinates of Projection* Points" *International Journal of Control, Automation, and Systems 8(5):, pp. 975-984*, 2010.

[8] N. Otsu, "A Threshold Selection Method from Gray Level Histograms". *IEEE Transactions on Systems, Man, And Cybernetics, vol. 9, pp.62-66*, 1979.

[9] Leonard G. C. Hamey, Colin Priest "Automatic Number Plate Recognition for Australian Conditions" *Proceedings of the Digital Imaging Computing Techniques And Applications(DICTA)* 2005

[10] Shishir Kumar, Shashank Agrawal and Kumar Saurabh "License Plate recognition system for indian vehicles" *International Journal of Information Technologyand Knowledge Volume 1 no.2 Management July*-December 2008

[11] Hanchuan Peng, Fuhui Long, and Zheru Chi, "Document Image Recognition Based on Template Matching of Component Block Projections" *IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 25, No. 9,*September 2003

[12] M.I. Khalil. "Car plate Recognition Using Template matching Method" *International Journal of Computer Theory and Engineering, Vol. 2, No. 5,* October, 2010

[13] Rule no.50,51,76 and 77, *Indian Central Motor Vehicle Act*,1989

[14] George Nagy "At the Frontiers of OCR". *Proceedings of the IEEE,vol.80 no.7,* July 1992

[15] Kumar Parasuraman, P.Vasantha Kumar "An Efficient Method for Indian Vehicle License Plate Extraction and Character Segmentation" *IEEE International Conference on Computational Intelligence and Computing Research, 2010.*

[16] Gonzalez R. C., Woods R. E., and Eddins S. L., "Digital Image Processing ", 3rd edition, Pearson Education(Singapore) Pvt. Ltd, 2009.