A Novel Approach for Recognition of Tamil Characters in Vehicle Number Plate based on Region Pixel through Surveillance Camera

E.K.Vellingiriraj, P.Balasubramanie

Abstract: ***Character segmentation is an important step in License Plate Recognition (LPR) system. There are many difficulties in this step, such as the influence of image noise, plate frame, rivet, the space mark, and so on. In natural language Number plate, there are additional problems faced to recognize the characters including Tamil characters. This work proposes a novel algorithm for character segmentation of degraded license plate based on Region pixel. Firstly, this algorithm performs preprocessing on the license plate, such as size normalization, uneven illumination correction, contrast enhancement, incline correction and edge enhancement; then, locates the character segments according to the vertical projection and merges the character segments that belong to the same character or splits the wider character segment according to the prior knowledge; finally, segments the characters according to the number and the width of the character segments. And last step convert the Tamil characters to equivalent English alphabets. This process is performed by the algorithm which takes the digital image as an input and gives textual form of license plate characters as an output. There are several methods used to perform this process.***

Keywords- (LPR).

I. INTRODUCTION

During the past few years, intelligent transportation systems (ITS) have had a wide impact on people's life as their scope is to improve transportation safety and mobility and to enhance productivity through the use of advanced technologies. License plate recognition (LPR) system is an important technique in ITS systems. There are many useful applications for a LPR (License Plate Recognition) system. The LPR algorithm consists of five steps: preprocessing, license plate locating, character segmentation, character recognition and character conversion. This paper presents a new algorithm for character segmentation. There are many factors that cause the character segmentation task difficult, such as image noise, plate frame, rivet, space mark, and plate rotation and illumination variance.

Character recognition is a broad term and consequently its application is possible in different areas and in different ways. In license plate character recognition, the complexity of the process enables the use of different methods or even combinations of methods to solve the problem. Unpredictable nature of input data (digital images) makes it difficult to choose a universal method that would be appropriate for every input image. Often the advantages of a method in one area are also disadvantages in the other.

Manuscript received August, 2013.

E.K.Vellingiriraj, Assistant Professor Department of Computer Science Engineering Kongu Engineering College Perundurai Erode – 638 052 TN India.

Dr.P.Balasubramanie, Professor Department of Computer Science Engineering Kongu Engineering College Perundurai Erode – 638 052 TN India.

Automatic license plate recognition is a process with the main aim of getting textual form of license plate characters from digital image of a vehicle. Digital image as an input passes through several stages of processing to obtain the string of characters as a final result. The core of this process is algorithm for optical character recognition (OCR) [1]. Algorithms like those are increasingly used in technical systems in transport. The most common examples of its usage are traffic monitoring, police investigations and identification on different kinds of entrances [2].

II. METHOD

This method consists of five steps. These steps are: preprocessing of an image, license plate detection, character segmentation, character recognition and character conversion.

2.1 Preprocessing:

The first step refers to processing and preparation of an image which is necessary for further license plate detection and character recognition. Pre-processing involves the application of digital filters to an image. Every colour image is first converted into a grayscale image to preserve memory and speed up further processing. This does not affect the useful data from the image. Threshold is used to accentuate license plate area. This is performed with an experimentally determined threshold value 170.After this step every pixel with a value larger than 170 becomes white (gets value 255) and every pixel with value smaller than 170 becomes black (gets value 0). As consequence, the grayscale image converts into a binary image with only two possible pixel values, black and white. Since the license plate is made of two contrasting colours, characters on it will remain visible [3].

Fig 1 (a) Colour Image

(b) Grayscale Image

Fig (1) Different type of threshold applied in image captured from surveillance camera.

Fig (1) shows a different type of threshold is applied in order to improve license plate area detection. Our idea is to count transitions from black to white pixels and conversely in each pixel row of the image. The array of the same pixel values must be at least 3 pixels long before a transition to avoid error caused by impurity of the input image. By using the number of these transitions, it is possible to exclude unnecessary parts of the image. If every license plate has at least 10 characters, there will be at least 20 transitions in rows within the license plate area. Therefore, each row with less than 20 transitions is not part of the plate area and contains unnecessary information. The purpose of additional thresholding is to black out these rows with useless data for easier further detection. Fig. 1 shows the steps in preprocessing of an image.

2.2 System Structure Chart:

Based on the system framework in the previous section, the Tamil Vehicle registration number image is converted into Tamil text format. This framework includes, i) Image scanning ii) Image preprocessing iii) Feature extraction iv) Character recognition v) Text conversion.

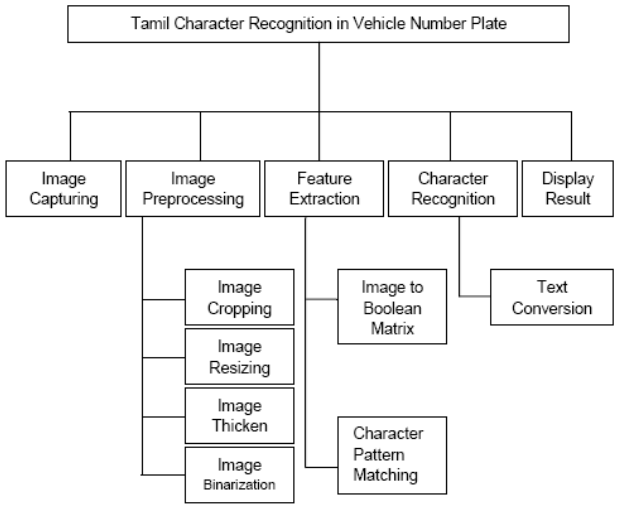

Fig (2) Structure chart of Tamil letters recognition from Vehicle Registration Plate by image zoning using the oolean matrix

i. *Image Capturing:* In the first stage, the tamil character capturing from vehicle number plate through surveillance camera and convert to the high resolution image.

ii. *Image Preprocessing:* In the image preprocessing module, the proposed system would prepare a character image for the feature extraction module. This stage consists of five sub-processes: a) image cropping b) segmentation c) image re-sizing d) image thickening and e) Image binarization. Each of these sub-processes are given below:

a) Image cropping: This process involves the cropping of each word. The image would have white space. Using this white space, words are cropped.

Fig (3) Character Image cropped from vehicle image

b) Segmentation: There are three types of segmentation. Line segmentation, word segmentation and character segmentation. Researchers used various techniques for segmentation like threshold techniques, region based method, edge based method, graph based methods, clustering methods, compression based methods, histogram methods, watershed transformation and model based methods.

Here we will take the edge detection method to segment the characters. The edge detection is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image brightness changes sharply or, more formally has discontinuities. Marr-Hildreth algorithm is one of the edge detection, which has a method of detecting edges in digital images.

c) Image re-sizing: Each segmented character is in different size. So, it is necessary to change all the characters into equal size. In the proposed method, the character image is re-sized into 100X100 pixels.

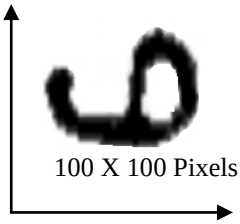

Fig (4) Character Image resized

d) Image thickening: Each darkened pixel of the re-sized character is thickened through darkening the nearest pixels. Using the nearest algorithm, a thin character is changed into thicker character through darkening the color of the nearest pixel for a target range.

Fig (5) Character image thicken

e) Image binarization: Each character is stored in Boolean matrix in either 0's or 1's. Using the image zoning technique [5], all the dark pixels are stored in 1's and light pixel in 0's.

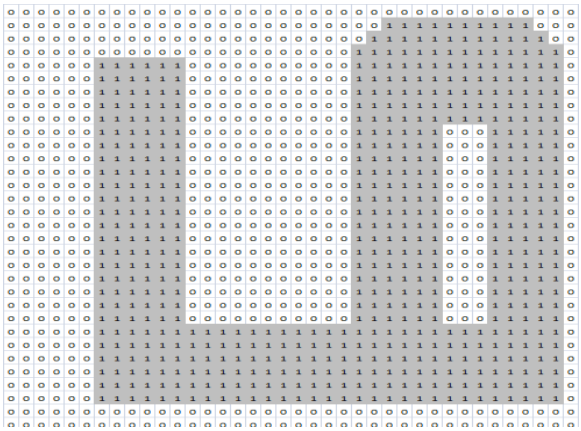

Fig (6) Character Image after binarization

iii. *Feature extraction:* This feature extraction module extracts the basic components of Tamil characters. There are three sub modules a) Image conversion into Boolean Matrix b) Image Grouping and c) Character Pattern Matching the details of which are given below:

a) *Image conversion into Boolean Matrix:* Each character from number plate is converted into Boolean matrix. Similarly, all the actual Tamil character sets (Fig 7.a & Fig 7.b) are also converted into the Boolean matrix.

Fig (7.a) Natural Numbers and equivalent Tamil Numbers

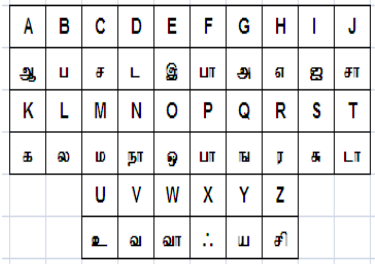

Fig (7.b) Alphabets and equivalent Tamil Characters

b) *Character Pattern Matching:* The stored original text Boolean matrix would be compared with the new preprocessing Tamil character using a range.

iv. Character Recognition: Pattern matching identifies each similar character in the predefined Boolean matrix. Then, the matching Boolean matrix is converted into equal Unicode Tamil fonts.

Fig 8 Segmentation for Vehicle number

2.3 Character Conversion

After recognize the characters in the number plate, each character convert to equivalent English characters like the State TN, District No XX, Serial No XX and the Registration No XXXX. The Tamil equivalent English data sets are stored predefined. In the Fig 7.a and Fig 7.b given the Tamil characters and numbers equivalent numbers and characters in English.

III. ALGORITHM IMPLEMENTATION

A graphic user interface is built for easier program testing through each phase from the previous section. Each phase can be performed by clicking a separate button, while the whole process can be performed by clicking one single button. Images, obtained using this algorithm, are shown one after another in the image-boxes. In the end, the recognized license plate characters are presented as a string inside the label in the lower right corner. Use of several integrated bitmap functions makes the work with images easier [7]. The functions are used for managing the pixel values and image cropping. The more complex functions described in previous section are coded and applied to image or part of image as needed.

A simple algorithm that is can be used for character matching is as follows:

M is the Boolean matrix one by one in all character sets in Tamil scripts. N is the Boolean matrix for current character in Number plate.

1. Predefine M in all Boolean matrix
2. Check the current character Boolean matrix N with M
3. If the range of both Boolean matrix is set to 1, the characters are identified
4. If not, check the part of the Boolean matrix with the entire predefined matrix
5. If Boolean matrix is matches, then check all the other parts of the Boolean matrix

IV. EXPERIMENTS

Experiments on a set of different images are performed to show the effectiveness of the implemented algorithm and used methods. The previously mentioned graphic user interface allows us to see intermediate results in each phase of the recognition process.

One example of the entire recognition process is described in detail in the previous method explanation. The following example shows the recognition process results for another loaded image. This test image shows a vehicle with a license plate in the natural environment. The proposed algorithm should be able to display license plate characters in a textual form as a final result, regardless of the bad factors (bright colours of the vehicle, impurity, background information, sun reflection). After gray scaling and first thresholding, the license plate area becomes prominent but not enough to isolate it. Additional threshold with a transition counter allows us to discard unnecessary information like white surfaces caused by sun reflection. The license plate area is then isolated correctly. Impurity between characters makes some noise problems which are solved by median filtering. Isolated license plate is segmented into individual characters. Finally, a correct string of license plate characters is obtained by the recognition algorithm. Expected accuracy of implemented method is experimentally obtained with 50 instances of input images. Test images were taken in different conditions and show the different vehicles. Correct recognition of license plates appeared in approximately 80 % of cases and about half of incorrect recognitions has the lapse in only one character.

V. CONCLUSION

It is difficult to achieve a robust algorithm that would be good in all segments and all cases of character recognition without fault. However, a combination of different methods and maximum utilization of computer possibilities in the processing of digital images allow us to bring the performance of character recognition on a high level. The algorithm implemented in this paper, except for a few common approaches in image processing, uses a combination of new ideas in detection and character recognition. These ideas refer to additional thresholding to make plate detection easier and use of characteristic thirds and ninths as a method for structural matching of characters. The experimental results show that the algorithm successfully recognizes characters in cases when there is no major angle distortion. Also, the algorithm avoids errors caused by noise or dirt on the license plate by applying an additional median filter. Although the recognition algorithm is based on structural characteristics of the characters, the font of characters may adversely affect the recognition result in case it varies from the font used on Croatian license plates which was the basis for writing the algorithm. Under normal conditions, the algorithm performs its tasks successfully, but because of several shortcomings, there is still room for

improvement and further development. This applies primarily to higher tolerance for input image deformation and increase of the number of recognizable font types.

REFERENCES

[1] Gonzalez, R. C.; Woods, R. E. Digital Image Processing, Prentice Hall, NewJersey, 2008.

[2] The Automatic Number Plate Recognition Tutorial. Quercus Technologies. 2006. URL: http://www.anpr-tutorial.com (15.08.2006.)

[3] Burger, W.; Burge, M. J. Digital Image Processing, Springer, NewYork, 2008.

[4] Fan, C. H.; Peng, Y. H. Vehicle License Plate Recognition System Design // Chung Hua Journal of Science and Engineering, 7, 2 (2009), 47-52.

[5] Burger, W.; Burge, M. J. Digital Image Processing, Springer, NewYork, 2008.

[6] Image processing learning resources, Morphology – Thinning. Hypermedia Image Processing Reference. 2003. URL:http://homepages.inf.ed.ac.uk/rbf/HIPR2/ thin.htm (2003.)

[7] Microsoft Developer Network Library - Bitmap Functions. Microsoft. 2012. URL: http://msdn.microsoft.com/enus/library/dd183385(v=vs.85).aspx (07.03.2012.)

[8] Y. Lu, "Machine printed character segmentation-an overview," Pattern Recognition, Publisher, vol. 28, no.1, pp. 67-80, 1995.

[9] R.G. Casey, and E. Lecolinet, "A survey of methods and strategies in character segmentation," IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.18, no.7, pp. 690-706, Jul.1996.

[10] S. Nomura, K. Yamanaka, O. Katai, H. Kawakami, and T. Shiose, "A novel adaptive morphological approach for degraded character image segmentation," Pattern Recognition, vol.38, no.11, pp. 1961-1975, Nov.2005.

[11] D. Michalopoulos, and C. Hu, "An error back-propagation artificial neural networks application in automatic car license plate recognition," Proceedings of the 15th international Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, Cairns, pp. 1-8, 2002.

[12] K.K.Kim, K.I.Kim, J.B.Kim, and H.J.Kim, "learning-based approach for license plate recognition," in Proceedings of the IEEE International Workshop on Neural Networks for Signal Processing, vol.2, pp. 614-623,2000.Giuseppe Pirlo, Donato Impedovo, "Adaptive Membership Functions for Handwritten Character Recognition by Voronoi-Based Image Zoning", IEEE Trans on Image Processing, Vol 21, No 9, PP 3827-3836, Sep 2012