Malaysian Vehicle License Plate Recognition

Othman Khalifa, Sheroz Khan, Rafiqul Islam, and Ahmad Suleiman
Kulliyyah of Engineering, International Islamic University Malaysia, Malaysia

Abstract: *Vehicle license plate recognition is an image-processing technology used to identify vehicles by their license plates. This technology can be used in various security and traffic applications, such as finding stolen cars, controlling access to car parks and gathering traffic flow statistics. In this paper an approach to license plate localization and recognition is presented. A proposed method to perform recognition of license plates under any environmental conditions, with no assumptions about the orientation of the plate or its distance from the camera is designed. To solve the problem of localization of a license plate, a simple texture-based approach based on edge information is used. Segmentation of characters is performed by using connected components analysis on license plate's image, and a simple multi-layer Perceptron neural network is used to recognize them. Simulation results were shown to be an efficient method for real time plate recognition.*

Keywords: *LPR, license plate, license plate recognition, OCR.*

Received March 14, 2006; accepted July 132006

1. Introduction

Vehicle License Plate Recognition (LPR) has become an important application in the transportation system. It can be used in many applications such as entrance admission, security, parking control, road traffic control, and speed control. A number of commercial software is developed in this area. However, they cannot be readily used when vehicle image is provided in different styles and formats. Also most of these software presume some constrains on the position and distance from the camera to vehicles, the inclined angles and the complexity of the captured image [5, 6, 12]. In this paper an approach is proposed to overcome these constrains and enhance the recognition process of the plate.

Image processing techniques such as edge detection, thresholding, resampling and filtering have been used to locate and isolate the license plate and the characters. Neural network is used for successful recognition of the license plate numbers. Once a license plate has been accurately identified, information about the vehicle can be obtained from various databases. The algorithm of License Plate Recognition (LPR) consists of the following steps:

1. Capturing the car's image.
2. Extracting the image of license plate.
3. Extracting characters from license plate image.
4. Finally, recognizing license plate characters.

This paper is organized as follows. A briefing of related work introduces in section 2. The proposed model and steps of the system are explained in section 3. Excremental results and simulation are presented in section 4. The conclusions and the further work are summarized in section 5.

2. Related Work

The problem of automatic vehicle license plate recognition has been studied since 1990s. The first approach was based on characteristics of boundary lines. The input image was first processed to enrich boundary lines information by some algorithms such as the gradient filter, and this resulted in an edging image. The image was binarized and then processed by certain algorithms, such as Hough Transform (HT), to detect lines. Eventually, couples of two parallel lines were considered as a plate-candidate [6]. However, boundary line detection is not suitable completely in the case of not horizontally recognizing the license plates on the image, corrupting or absentness of the boundary line in the license plates due to noise and uneven brightness. Furthermore, HT is inherently a time-consuming process.

The color and textures of the license plate have also been used to identify the license plate [9, 11], but they seem to be aimless and ineffective, especially when the system has plates of different colors and sign patterns.

Another common approach involves the use of Artificial Neural Networks (ANNs) [7].

3. Proposed Model

The main objective of this paper is to develop a system that can extract the license plate number from complex-scene images with no assumptions of image quality, the distance between the camera and the plate, the angle in which the plate has been captured relative to the camera and so on.

An overview of the proposed number-plate recognition system can be seen in Figure 1. After the vehicle image is captured by the camera, it will be passed to a pre-processing unit which prepares the image for further processing by the system. Its main operations are to eliminate the noise caused by the image-acquisition subsystem and enhance the image features used by the other two subsystems. The image will be scanned by the Plate Extraction module to locate the vehicle's license-plate. The next phase is responsible for segmenting the individual characters of the plate. Finally each character will be passed to the Optical Character Recognition (OCR) module to be identified and the final results are ASCII characters and numbers of the plate.

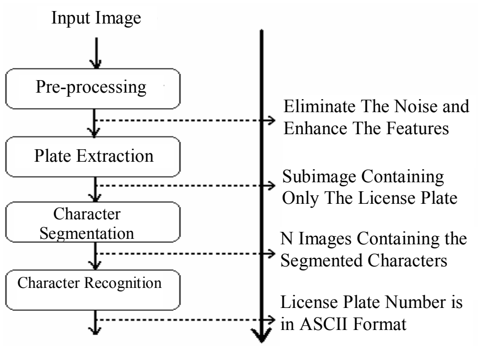


Figure 1. License plate recognition model.

3.1. Pre-Processing

The input image is initially processed to improve its quality and prepare it to next stages of the system. First, the system will convert RGB images to gray-level images using the NTSC standard method:

$$Gray = 0.299*Red + 0.587*Green + 0.114*Blue$$

In the second step, a median filter (5x5) is applied to the gray-level image in order to remove the noise, while preserving the sharpness of the image. The median filter is a non-linear filter, which replaces each pixel with a value obtained by computing the median of values of pixels -in this case- in a 5x5 neighborhood of the original pixel.

3.2. Plate Extraction

The plate extraction process contains five different phases as illustrated in Figure 2; here each phase performs a segmentation process on the gray image to eliminate the redundant pixels that don't belong to a plate region. For example, the horizontal localization phase is responsible of finding horizontal segments that may contain a license plate. In the following discussion, each of these phases will be discussed in details.

Malaysian license plates consist of a row of white characters on a black background, so we can say that the license plate region is characterized by a row of transitions from black to white and vice versa, such transitions are called "edges". The total change in intensity from plate's characters to its background is called the strength of the edge. The strongest edge value that can occur can be found in the case of a transition from a black pixel to a white pixel or from a white pixel to a black pixel. This is the ideal case in Malaysian license plates where the characters are white drawn on black background, and hence they produce high intensity edge values which can be used to find possible plate regions. In this paper we are going to use Sobel operators to find the edge image. The Sobel operator performs a 2-D spatial gradient measurement on an image. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image. The Sobel edge detector uses a pair of 3x3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows) [3]. The actual Sobel masks are shown below:

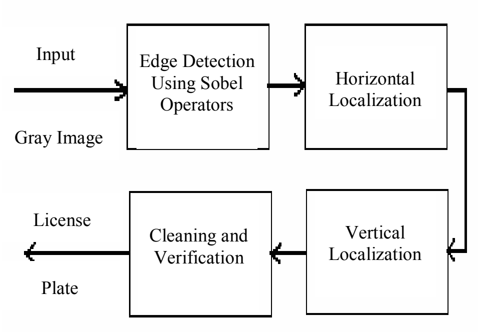


Figure 2. Plate extraction module.

-1	0	+1
-2	0	+2
-1	0	+1

(a)

+1	+2	+1
0	0	0
-1	-2	-1

(b)

Figure 3. Sobel masks for edge detection, (a) Vertical, (b) Horizontal.

Figure 4 illustrates an edge detection process using Sobel operator. The lighter region indicates stronger edges as the case for license plate.

After creating the edge image, the system will search for regions with high edge values which are most likely to contain a license plate. To do so, the system will construct a horizontal projection profile for the edge image. A projection profile of an image is a compact representation of the spatial pixel content distribution. A horizontal projection profile is defined as the vector of the sums of the pixel intensities over each row. Figure 5 shows a graph representation of the horizontal projection of the edge image, where the peaks of the graph indicate regions with strong edges and can be used to extract the most likely horizontal positions of the license plate. The system will label all the horizontal segments that correspond to significant peaks in the projection profile of the edge image.

(a) Original image

(b) Edge image

Figure 4. Sobel edge detection.

The next step in the plate extraction process is to find the vertical position of the license plate. For this part, the methods generally use a statistical study of histograms of the picture of the plate [3], but all these methods are not applicable in the case of Malaysian license plates for two reasons:

1. Number of characters varies from plate to another.
2. Some plates have a square shape with the characters written in two lines.

Figure 5. Horizontal projection of the edge image.

We propose a simple method to locate the vertical coordinates of the plate using the edge image. The algorithm starts by sliding a window through each horizontal segment and sums the values of edges inside the window and then divides the result by the window's area to get the average gradient magnitude per pixel inside the window at that specific location. For each horizontal segment, the algorithm will store the resulted value of each window step in a new vector for later processing, we shall call this vector an HDV and can be calculated as:

$$HDV(i) = \frac{\sum_{x \in W} \sum_{y \in W} G(x, y)}{A_w}$$

where, i = 0, 1… (ImW – M + 1), $G(x, y)$ is the edge image, W is the sliding window, Aw is the area of the sliding window, ImW is the width of the image and M is the width of the sliding window.

Figure 6 illustrates the computation process of the horizontal density vector. As the window slides over the plate the value of the Horizontal Density Vector (HDV) graph increases gradually until it reaches the peak of the graph and then decreases again, at the peak point in the HDV graph the sliding window aligns with the license plate.

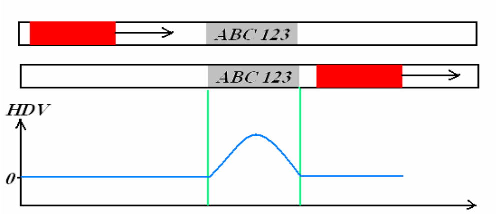
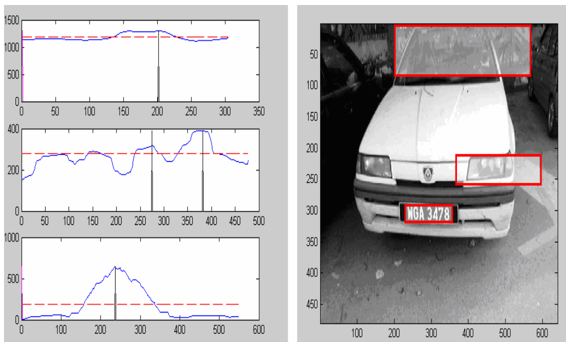


Figure 6. Computation of HDV.

After that, the algorithm will find all the peaks in the HDV graph for each horizontal segment. A peak is assumed to belong to a license plate if it is greater than a determined threshold. Figure 7 shows an example for the process of extracting the vertical position of the license plate along with the HDV graph for each horizontal segment.



(a) (b)

Figure 7. Example of vertical localization of the license plate.

The next step after the vertical localization process is the cleaning phase. This phase performs three main processes which are:

1. The segmentation of two-row plates which were wrongly detected as one segment.
2. The isolation of the license plate from any redundant background.
3. Skew angle detection and correction.

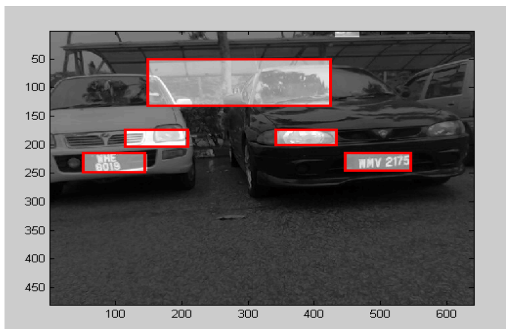


Figure 8. Two plates appear in the same horizontal segment.

The segmentation of plates that contain two rows of characters will be done through a horizontal projection of the binary image for the extracted segment and searching for a valley that goes beyond a specific threshold. If found, the segment will be considered as a two row plate and will be cut into two separate segments, as shown in Figure 9 below:

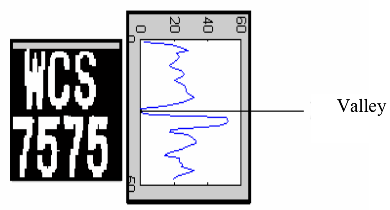


Figure 9. Horizontal projection of the plate.

The isolation of the license plate from any superfluous background is performed using a projection profile that reflects the number of white pixels in each row and in each column. In most cases, projecting the amount of white pixels both vertically and horizontally reveals the actual position of the plate. An example of the projections is shown in Figure 10.

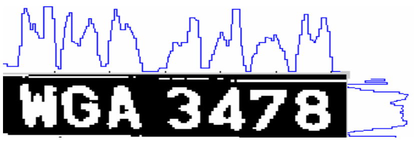


Figure 10 Horizontal and vertical projection of the license plate.

To find the skew angle of the plate we used the algorithm described in [1]. The computation of the skew angle is achieved using Hough transform to find the most visible lines in the plate image and their angles, and finally the plate image will be rotated as shown in Figure 11.

Figure 11 Skew angle correction.

The last step in the localization process is the verification phase. The input to this phase is all the extracted regions from the previous stages which may contain license plates. So the main objective of this phase is to filter out the regions that do not belong to a license plate. We will follow a rule-based classification approach, where an extracted region will be deleted if one of the following conditions is satisfied:

1. Aspect ratio (width/height) of the region is smaller/greater than a threshold.
2. The area of the region is less than a threshold value.
3. Average gradient magnitude per pixel inside the region is smaller/greater than a threshold value.
4. The number of valid characters is less than 3 or greater than 9.
5. Density of the region is less/greater than a threshold *(density = objects area ÷ total area)*. All thresholds are determined experimentally from the available training data.

3.3. Character Segmentation

To ease the process of identifying the characters, it is preferable to divide the extracted plate into different images, each containing one isolated character. There are some widely used methods for character isolation which are used in almost all available LPR systems. Those methods are: static bounds [4], vertical projection [8] and Connected-Components. The first two methods can not be used to segment Malaysian license plates since they do not have a fixed number of characters for each plate. In this paper the following steps are used to segment the characters of the license plate:

1. Stretch the contrast of the image to extend over the entire range of gray levels available (0-255).
2. Threshold the plate image using Otsu method [10].
3. Search for connected components in the image, each connected component will be assigned a special label in order to distinguish between different connected components in the image as shown in Figure 12.
4. Resize each character from the previous step to the standard height and width (20x10) in order to be used in the following recognition process.

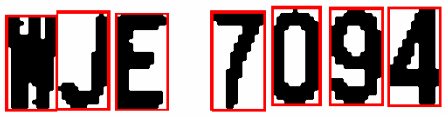


Figure 12 Extracting the characters from license plate.

3.4. Character Recognition

ANN is used for character recognition. We used a Multi-Layer Perceptron Neural Network (MLP NN) trained with the back-propagation algorithm. During learning phase, characters of the constituted database are successively presented at the input layer of MLP network and their corresponding outputs are compared to the desired outputs. Weights are iteratively modified in order to obtain, at the network outputs, responses which are as close as possible to the desired outputs. The input layer consists of 135 input neurons and there is one hidden layer which has 40 neurons with log-sigmoid transfer functions. The output stage has 36 neurons, each outputting the probability that the given input is the corresponding character.

4. Experimental Results

Experiments have been performed to test the proposed system. The system is simulated in MATLAB6.5.1 for the recognition of Malaysian license plates. A set of 150 images with various sizes were used for testing. The images were taken from:

1. Complex scenes, in which several objects with complex textures are presented.
2. Various environments (street, roadside and parking lot).
3. Different inclined angles and distances relative to the camera.

Performance for individual system subsections were as follows: 92.1% of plates were located successfully; characters were segmented correctly from 90.5% of manually extracted plates, 93.2% of these characters were correctly classified.

The major source of errors in plate extraction process and character segmentation was mainly due to bad quality of input image during the acquisition stage. Figure 13 shows some simulation results for the proposed system.

Figure 13. Simulation results.

5. Conclusions

The purpose of this paper is to investigate the possibility of making a system for automatic recognition of license plates. Given an input image, it should be able to firstly extract the license plate, then isolate the characters contained in the plate, and finally identify the characters in the license plate. The proposed system will search the image for high density edge regions which may contain a license plate. After that a cleaning and a verification processes will be performed on the extracted regions to filter out those regions that do not contain a license plate. After that the plate will be passed to the segmentation phase where it will be divided into a number of sub-images equal to the number of characters contained in the plate. Finally, each of the sub-images will be passed to a Multi-Layer Perceptron Neural Network for identification.

References

[1] Farrow G. S. D., Ireton M. A., and Xydeas C. S., "Detecting the Skew Angle in Document Images," *Signal Processing: Image Communication*, vol. 6, no. 2, pp. 101-114, 1994.

[2] Fredric M. and Kostanic I., *Principles of Neurocomputing for Science and Engineering*, McGraw Hill, 2000.

[3] Gonzalez R. C. and Woods R. E., *Digital Image Processing*, Addison-Wesley, 1993.

[4] http://en.wikipedia.org/wiki/Automatic_number_plate__recognition, Retrieved, August 2005.

[5] http://www.licenseplaterecognition.com, August 2005.

[6] Kamat V. and Ganesan S., "An Efficient Implementation of Hough Transform for Detecting Vehicle License Plate Using DSP'S," *in Proceedings of the Cellular Neural Networks and their Applications, Proceedings the IEEE 31st Annual*, pp. 209-218, 1997.

[7] Kim K. K., Kim K. I., Kim J. B., and Kim H. J., "Learning Based Approach for License Plate Recognition," *in Proceedings of IEEE Processing Society Workshop on Neural*

Networks for Signal Processing, vol. 2, pp. 614-623, 2000.

[8] Lu Y., "Machine Printed Character Segmentation," *Pattern Recognition*, vol. 28, no. 1, pp. 67-80, Elsevier Science Ltd, UK, 1995.

[9] Nijhuis J. A. G., Brugge M. H., and Helmholt K. A., "License Plate Recognition Using DTCNNs," *in Proceedings of the 5th IEEE International Workshop on Publish*, London, pp. 212-217, 1998.

[10] Otsu N., "A Threshold Selection Method for Gray Level Histograms," *IEEE Transactions on System, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, January 1979.

[11] Parker J. R. and Federl P., "An Approach to Licence Plate Recognition," *Technical Report*, The Laboratory for Computer Vision, University of Calgary, 1996.

[12] Setchell C. J., "Applications of Computer Vision to Road-traffic Monitoring," *PhD Thesis*, Bristol University, Bristol, UK, September 1997.

Othman Khalifa received his Bachelor degree in electronic engineering from the Garyounis University, Libya in 1986, his Master degree in electronics science engineering, and his PhD in digital image processing from Newcastle University, UK in 1996 and 2000, respectively. He worked in industrial for eight years. Currently, he is head of the Department of Electrical and Computer Engineering, International Islamic University Malaysia. He has 65 publications in international journals and conferences. His area of interests includes communications, information theory and coding, digital image, video processing, coding and compression, wavelets, fractal and pattern recognition. He is IEEE member, Image Processing and Communication Society member

Sheroz Khan received his BSc in electrical engineering from NWFP University of Engineering and Technology, Peshawar, Pakistan in 1982, his MSc in micro-electronic & computer engineering from Surrey, Guildford, UK in 1990 and his PhD in AI applications to SCADA systems from the University of Strathclyde, Glasgow, UK in 1994. He has been actively involved in the teaching and research activities at the university level in Peshawar and Malaysia. He joined University Tenaga National (UNITEN) January 2000. Currently, he is an associate professor within the Department of Electrical and Computer Engineering, IIUM. He has authored/co-authored more than 35 publications in both local and international journals and conference proceedings. His area of interests includes data acquisition, embedded systems, interfacing and intelligent control system, design and instrumentation. He is a member of Pakistan Engineering Council and a member of the IEEE Instrumentation and Measurement Society.

Rafiqul Islam received his BSc in electrical and electronic engineering from BUET, Dhaka in 1987, his MSc and PhD both in electrical engineering from UTM in 1996 and 2000, respectively. He is fellow of IEB and member of IEEE. Currently, he is faculty member of Electrical and Computer Engineering Department of International Islamic University Malaysia. His area of interests includes radio link design, RF propagation measurement and RF design, smart antennas, and array antennas design.

Ahmad Suleiman is a graduate student, International Islamic University Malaysia, 2006. His area of interests includes signal/Image processing and digital communications.