A NEW PROPOSED METHOD FOR AUTOMATIC NUMBER PLATE RECOGNITION

Bui Huu Phu[1], Trinh Hoang Hon[2]

[1]*Institute of Applied Mechanics and Informatics; huuphu@iami.vast.vn*
[2]*Hochiminh City University of Technology; trinhhoanghon09@gmail.com*

Abstract - This paper describes our new proposed method for automatically identifying the text of vehicle number plate. From a new image, the method automatically extracts plate, classifies and recognizes characters (digits). Firstly, Laplace operator is used to calculate the magnitude of gradient which then is binarized based on a threshold. In the second phase, a sub-window slides all images to search the candidate of plate.The best candidate is chosen by aspect ratio conditions. The extracted plate is then enhanced and classified by colour segment conditions to find candidates of characters as foreground. The horizontal and vertical histograms are used to classify and extract each character separately. Finally, Support Vector Machine algorithm is used to identify the characters. The experimental results have shown that our proposed method obtains high accuracy and is available for real applications.

Key words - Automatic number plate recognition (ANPR); Laplace operator; plate extraction; plate classification; plate identification.

1. Introduction

Automatic number plate recognition (ANPR), also called other names as Automatic license-plate recognition (ALPR), Automatic license-plate reader (ALPR), Automatic vehicle identification (AVI), Car plate recognition (CPR), License-plate recognition (LPR), Lecture automatique de plaques d'immatriculation (LAPI), Mobile license-plate reader (MLPR), or Vehicle license-plate recognition (VLPR), is a technology that uses optical character recognition on images to read vehicle registration plates. It can use existing closed-circuit television, road-rule enforcement cameras, or cameras specifically designed for the task. ANPR is used by police forces around the world for law enforcement purposes. It is also used for electronic toll collection on pay-per-use roads and as a method of cataloging the movements of traffic for example by highway agencies.

Automatic number plate recognition can be used to store the images captured by the cameras as well as the text from the license plate, with some configurable to store a photograph of the driver. Systems commonly use infrared lighting to allow the camera to take the picture at any time of the day [1-3]. ANPR technology tends to be region-specific, owing to plate variation from place to place.

ANPR was invented in 1976 at the Police Scientific Development Branch in the UK. Prototype systems were working by 1979, and contracts were let to produce industrial systems, first at EMI Electronics, and then at Computer Recognition Systems (CRS) in Wokingham, UK. Early trial systems were deployed on the A1 road and at the Dartford Tunnel. However it did not become widely used until new developments in cheaper and easier to use software was pioneered during the 1990s. The first arrest through detection of a stolen car was made in 1981 and the first documented case of ANPR in helping solve a murder occurred in November 2005 after the murder of Sharon Beshenivsky, in which City of Bradford based ANPR played a vital role in locating and subsequently convicting her killers [4].

The software aspect of the system runs on standard home computer hardware and can be linked to other applications or databases. It first uses a series of image manipulation techniques to detect, normalize and enhance the image of the number plate, and then optical character recognition (OCR) to extract the alpha-numeric of the license plate. ANPR systems are generally deployed in one of two basic approaches: one allows for the entire process to be performed at the lane location in real-time, and the other transmits all the images from many lanes to a remote computer location and performs the OCR process there at some later point in time. When done at the lane site, the information captured of the plate alphanumeric, date-time, lane identification, and any other information required is completed in approximately 250 milliseconds. This information can easily be transmitted to a remote computer for further processing if necessary, or stored at the lane for later retrieval. In the other arrangement, there are typically large numbers of PCs used in a server farm to handle high workloads, such as those found in the London congestion charge project. Often in such systems, there is a requirement to forward images to the remote server, and this can require larger bandwidth transmission media.

ANPR uses optical character recognition (OCR) on images taken by cameras. When Dutch vehicle registration plates switched to a different style in 2002, one of the changes made was to the font, introducing small gaps in some letters (such as P and R) to make them more distinct and therefore more legible to such systems. Some license plate arrangements use variations in font sizes and positioning—ANPR systems must be able to cope with such differences in order to be truly effective. More complicated systems can cope with international variants, though many programs are individually tailored to each country.

There are seven primary algorithms that the software requires for identifying a license plate:

- Plate localization – responsible for finding and isolating the plate on the picture.

- Plate orientation and sizing – compensates for the skew of the plate and adjusts the dimensions to the required size.

- Normalization – adjusts the brightness and contrast of the image.

- Character segmentation – finds the individual characters on the plates.

- Optical character recognition.

- Syntactical/Geometrical analysis – check characters and positions against country-specific rules.

- The averaging of the recognized value over multiple fields/images to produce a more reliable or confident result. Any single image may contain a reflected light flare whether it is partially obscured or other temporary effect.

The complexity of each of these subsections of the program determines the accuracy of the system. During the third phase (normalization), some systems use edge detection techniques to increase the picture difference between the letters and the plate backing. A median filter may also be used to reduce the visual noise on the image.

ANPR has been widely studied in the world. In [6], the author has described a fast algorithm for automatic license plate detection system for the Egyptian license plates that achieves a high detection rate without the need for a high quality images from expensive hardware. The system captures images of the vehicles with a digital camera. An algorithm for the extraction of the license plate has been explained and designed using Matlab. The result can be achieved at about 96% detection rate for small dataset. In [7], as few constraints as possible on the working environment are considered. The proposed LPR technique consists of two main modules: a license plate locating module and a license number identification module. The former characterized by fuzzy disciplines attempts to extract license plates from an input image, while the latter conceptualized in terms of neural subjects aims to identify the number present in a license plate. Experiments have been conducted for the respective modules. In the experiment on locating license plates, 1088 images taken from various scenes and under different conditions were employed. Among them, 23 images have been failed to locate the license plates present in the images; the license plate location rate of success is 97.9%. In the experiment on identifying license number, 1065 images, from which license plates have been successfully located, were used. Among them, 47 images have failed to identify the numbers of the license plates located in the images; the identification rate of success is 95.6%. Combining the above two rates, the overall rate of success for our LPR algorithm is 93.7%. In [8], the proposed method applied on yellow color license plate. It has two main stages. Firstly, exact location of the license plate is detected from an input image by using image acquisition and optical character recognition and Sobel edge is used for character segmentation. Secondly, template matching is used to test the recognized characters with templates. This paper also proposes vehicle authorization by checking the license plate number from database and electronic mail is send to administrator if authorization fails. In [9], the Edge Detection methods are used to locate the rectangles from an image. This is very simple and fast technique. Morphology [10] is used to extract the license plate from the original image. It helps to remove unwanted small parts from license plate.

In Vietnam, the demand for the software of ANPR is very high in many applications. Until now, in the Vietnam market, only the imported software licenses with the high price have been used. Therefore, the authors want to develop a domestic ANPR with the same quality but much cheaper than imported ones.

In the paper, the authors present our own proposal of ANPR. We have developed a new algorithm, then build the software, and test it in the realistic environment. The results have shown that our ANPR software has worked very well and given high precision nearly 100% under the industrial test. The authors believe the software can be packed to be commercialized in the near future.

The remaining of the paper has the structure as follows. In section 2 the model of ANPR is presented. In sections 3 and 4, our proposed algorithm and processing steps of ANPR are described. The experimental results are shown in section 5. Finally, the conclusions are discussed in section 6.

2. Model of Automatic Number Plate Recognition

In general, the processing steps of ANPR are shown in Figure 1. The input images got from digital cameras will detect and extract the plate region, and then the software will classify and identify characters and digits. Finally it gives the result of recognizing the number plate of the vehicle.

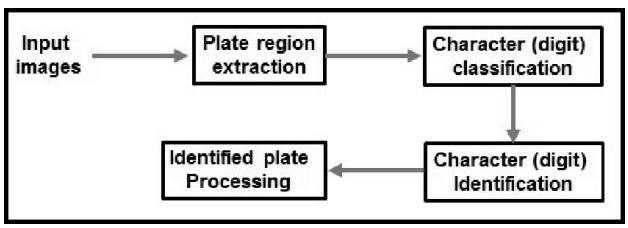


Figure 1. *Proposed scheme for identifying plate vehicles*

In ANPR, the steps of plate extraction, character (digit) classification, and character (digit) identification are the most important, and decide the quality and correction of the software.

a. one row of characters/digits b. two rows of characters/digits

Figure 2. *Several types of vehicle plates in Vietnam*

Because the number plates of vehicles are different in countries, the processing algorithms will be different. In Vietnam, basically, there are 2 common kinds of vehicle plates shown in Figure 2. The number plate recognition of vehicles in Vietnam is quite difficult because many number plates are not clear, bended, and mounted with strange things. In the following sections, our proposed methods in the important steps of ANPR are presented.

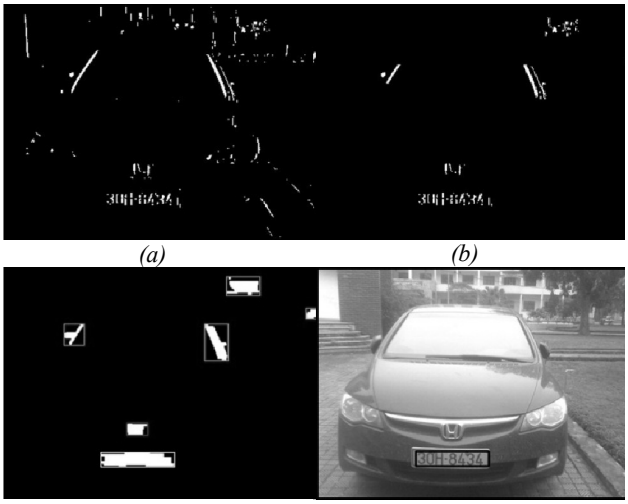
3. Plate Extraction

The vehicle plate structure in Vietnam is determined by the Vietnamese government. The dimension and font size of vehicle plates are standardized commonly. There are usually two kinds of plates as one row or two rows characters, shown in Figure 2. The dimension of one row character plates is 470mmx110mm, the height of characters is approximately 75% but not higher than 80% of the plates, and the width of characters equals a half of the height. For the plates with two row characters, the dimension is 280mmx200mm, but the size of characters is similar to the first kind.

In the realistic environments, the plates are affected by the distortion such as scale, skew, rotation, and cameras. So in the process, the relative dimensions are used with tolerance ratio between plate's dimensions built by the experience.

(a) *(b)*

Figure 3. *(a) Original image, (b) Magnitude of gradient with Laplace operation*



(a) *(b)*

(c) *(d)*

Figure 4. *Plate extraction; (a) Binary image, (b) survived image, (c) aspect ratio verification, (d) example of results of plate extraction.*

Firstly, the input image (Figure3 (a)) is performed to get edge enhancement by Gaussian operator [11] as Eq.1 whose σ is chosen $\sqrt{2}$; and Gaussian Kernel size is 3 × 3 pixels.

$$G(x,y) = \frac{1}{\sqrt[2]{2\pi\sigma^2}}\exp(-\frac{x^2+y^2}{2\sigma^2}), \quad (1)$$

The results (an example in Figure3 (b)) are binarized with the threshold of $T = 0.07$, then a subwindow with 16×8 size is slided to search candidates of plates, the candidates which satisfy condition in (2) will survive.

$$\sum_{i=1}^{h}\sum_{j=1}^{w} BW_{(i,j)} \geq 15, \quad (2)$$

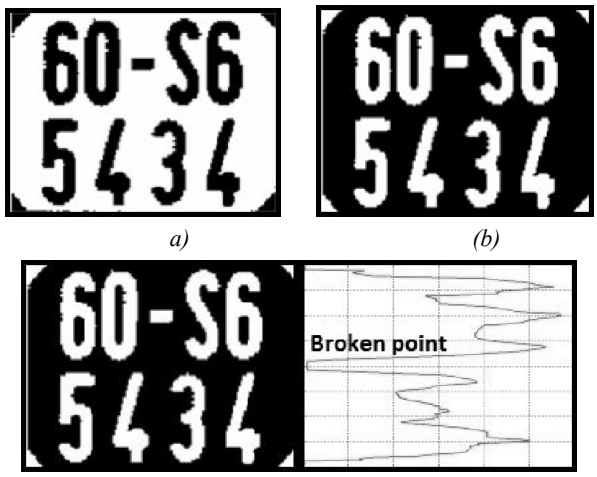
where, BW is considered as binary image, w and h are the width and height of sub-window. Threshold value is chosen as 15 by our experience. A demonstration of this step can be seen in Figure 4 (b).

A bounding box of each plate candidate is calculated, the candidate is chosen from candidates if its aspect ratio is passed the condition in the non-equation (3), whereas non-equations 3(a) and 3(b) are applied for one row and two row character plates respectively. Figure 4 (d) shows a result of plate detection.

$$\begin{cases} 0.15 \leq \frac{H}{W} \leq 0.3 \ \ for\ one\ row\ character\ plates \ \ (3a) \\ 0.5 \leq \frac{H}{W} \leq 0.95 \ \ for\ two\ row\ character\ plates \ \ (3b) \end{cases}$$

4. Character (Digit) Identification

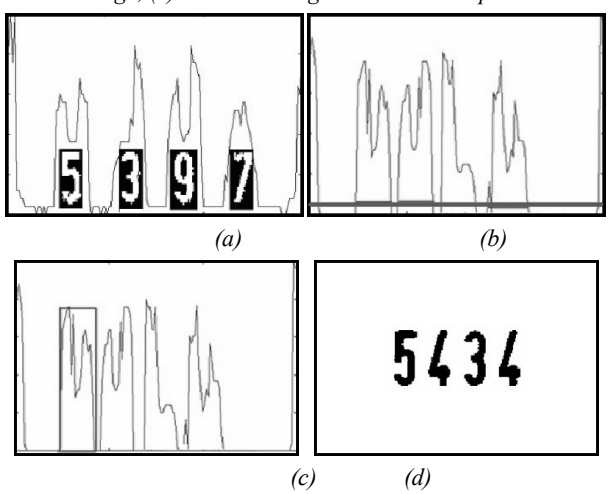
The extracted plates are binarized to get foreground shown in Figure 5a, the intensity of pixel is inverted to get the white foreground as shown in Figure 5b. Four line segments nearby bounding box are used to retrieve the rectangular shape. To do so, the 2D-2D projection is used. Then the plates are normalized to the real aspect ratio of $\frac{H}{W} = \frac{35}{150}$.



a) *(b)*

(c)

Figure 5. *(a) enhanced characters of two row plate, (b) inverted image, (c) vertical histogram and broken point*



(a) *(b)*

(c) *(d)*

Figure 6. *Character (digit) classification: (a) horizontal histogram, (b) characters separated, (c) width and area of the candidate of characters, (d) example of results of character classification*

In the case of two row character plates, a horizontal histogram is set up to separate the first and the second row of character. A vertical histogram of each row of both kinds of plates is calculated as shown in Figure 6. Two thresholds are issued for this contest, one is a character width (T_{CW}) and the other is the min area (A_{min}). (T_{CW}) is selected by from 70% to 130% the real size;. This means that

$$9 \le T_{\text{CW}} \le 17 \tag{4}$$

pixels. For example

$$\text{T}_{\text{CW}}^{\text{min}} = 0{,}7 \text{ x } 40 \text{ x } \frac{150}{470} = 8{,}9 \approx 9 \tag{5}$$

pixels, while A_{min} is defined as the product of the character width and the peak of vertical histogram. All the conditions are demonstrated in Figure 6.

In practice, if the horizontal histogram of two row character plates and/or the vertical histogram have not been broken points, this means that the rows or characters could not be classified, then the plates are treated as the negative error of the previous step; the negative error defines that non-plate is undetected as a plate.

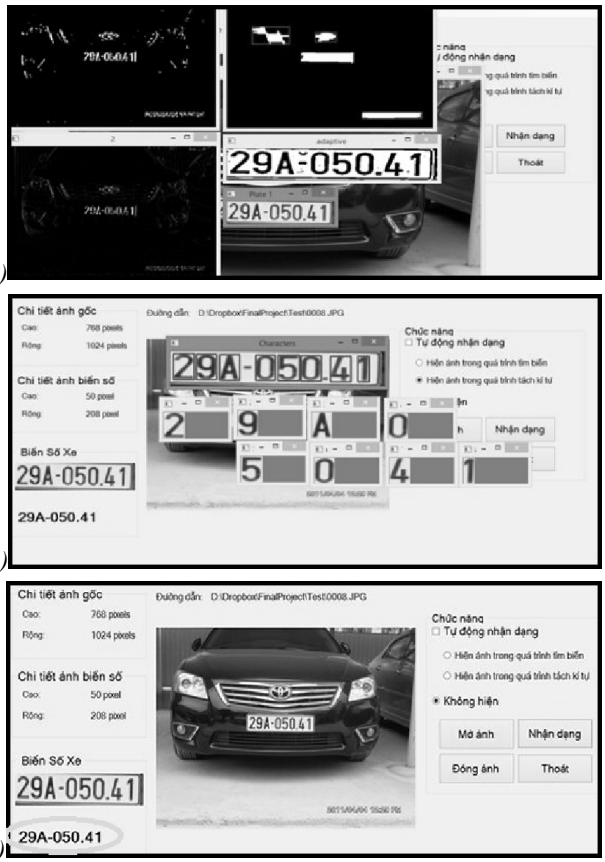


Figure 7. *Complete process steps in ANPR: (a) plate extraction, (b) character/digit classification, (c) character/digit identification*

The last step for this phase is character identification; the supported vector machine [1] algorithm is used to identify the extracted characters. The training set of characters is automatically extracted as previous steps, and then 50 correct ways of each character or digit are chosen by program. They are then used as the input data for the training function. The training process is performed in offline mode.

Figure 7. illustrates a complete process steps in ANPR, consisting of the input image and the result of plate extraction, shown in Figure 7(a); the result of character classification, shown in Figure 7(b); and the character identification, shown in Figure 7c.

5. Experimental Results and Discussions

Based on the proposed algorithms, the authors have developed the ANPR software completely. The software is developed based on C++ through Visual Studio 2008, and uses image processing library OpenCV version 2.4.9. The authors have set up a ANPR system at An Suong An Lac Tolling Station for testing the software and system. The camera used for testing our software is IRLab with the model CIR- HUW34WP and the resolution of 700TVL.

During the time testing, there are about 10000 vehicle samples collected and tested. The authors have evaluated that with about 9000 clear and good appearance plates, the ANPR system can give very good and correct results. The system can work well in day time and night time. In addition, the software can well recognize the number of the blue and red color plates. Below are some results that have done at An Suong An Lac Tolling Stations.

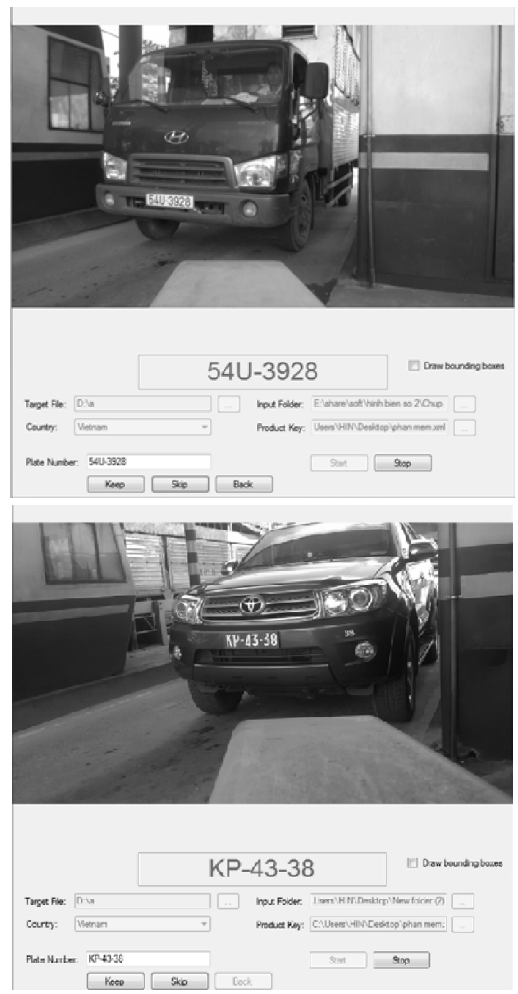


Figure 8. *Some test results of the proposed ANPR system*

Although, the ANPR software has given a good result, in some realistic cases, the system can not recognize the number of plates. After reviewing and evaluating, we see

that there are some reasons why the ANPR can not work well as follows:

- The plates are damaged, curved, distorted.
- The plates are attached to some abnormal subjects.
- The plates are too dim

In the future, we will consider developing new methods to improve the quality and correction of our ANPR system.

6. Conclusions

In the paper, we have presented our proposal to develop ANPR algorithm and system. The important steps in the proposal have been described in detail. After completely building the software and setting up the system at An Suong An Lac Tolling Collection Station, the authors have tested it under the realistic environments for a month with about 10000 vehicle number plate samples. As a result, we see that our proposed ANPR software can work very well and give very high precision of number recognition. In the standard condition, the ratio of the correct number recognition is approximately 100%. However, in some special cases when the number plates are damaged, distorted, or dim, the software does not work well.

With the result, it is obvious that our proposed and developed ANPR software can be packed to be used for realistic applications in the future.

Acknowledgment

The authors would like to thank National Key Laboratory of Digital Control and System Engineering (DCSELAB) and Hochiminh City Department of Science and Technology for supporting and funding the research project.

REFERENCES

[1] "ANPR Tutorial". *ANPR Tutorial. 15 August 2006*. Retrieved 2012-01-24.

[2] Shan Du ; IntelliView Technol., Inc., Calgary, AB, Canada ; Ibrahim, M. ; Shehata, M. ; Badawy, Wael; Automatic License Plate Recognition (ALPR): A State-of-the-Art Review". *IEEE. 1 Feb 2013*. Retrieved 2014-01-09.

[3] "An introduction to ANPR". *Cctv-information.co.uk*. Retrieved 2012-01-24.

[4] "CCTV network tracks getaway car". BBC News. 21 November 2005. Retrieved 2013-08-12.

[5] Amr E. Rashid, "A fast algorithm for license plate detection," International Conference on Signal Processing Image Processing & Pattern Recognition (ICSIPR), pp. 44-48, Feb. 2013.

[6] S.L. Chang, L.S. Chen, Y.C. Chung, and S.W. Chen, "Automatic license plate recognition," IEEE Transactions on Intelligent Transportation Systems, vol. 5, pp. 42-53, Mar. 2004.

[7] V. Sharma, P. C. Mathpal, and A. Kaushik "Automatic license plate recognition using optical character recognition and template matching on yellow color license plate", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, pp. 12984 – 12990, May 2014.

[8] Zheng, D., Zhao, Y., and Wang, J., "An Efficient Method of License Plate Location", Pattern Recognit. Lett.., vol.26, pp.2431-2438, 2005.

[9] Nelson, C., and Babu,,K., " A License Plate Localization using Morphology and Recognition", IEEE India conference, pp.34-39, 2008.

[10] Dingyun, W., Lihong, Z., and Yingbo, L., "A New Algorithm for License Plate Recognition Based on Improved Edge Detection and Mathematical Morphology", IEEE International Conference on Information Science and Engineering , pp.1724-1727, 2010.

[11] G.T. Shrivakshan, and C. Chandrasekar, "A Comparison of various Edge Detection Techniques used in Image Processing", International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012.

(The Board of Editors received the paper on 22/04/2016, its review was completed on 05/05/2016)