*Research Article*

# Thai License Plate Recognition System From a Video Stream

Chomtip Pornpanomchai* and Nuchakarn Anawatmongkon

*Faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand*
**Corresponding author E-mail address: chomtip.por@mahidol.ac.th*

Received December 21, 2013; Accepted April 8, 2014

**Abstract**

The objective of this project was to develop a computer system which is automatically capable of localization, segmentation and recognition of a Thai license plate from a video file. The system was called "Thai license plate recognition system" or "TLPRS". The TLPRS consisted of 5 modules, which were 1) video image acquisition, 2) license plate localization, 3) character segmentation, 4) character recognition, and 5) result presentation. The experiment was conducted on video files and localized car license plates from video frames in every second. After that the system segmented each character and recognized it. The experiment was conducted on video files with total length of 2.5 hours, which consisted of 497 car images. Based on the experimental results, the precision rate of license plate localization was 91.95 percent, while the character segmentation rate was 64.19 percent and the character recognition rate was 93.26 percent.

***Key Words:*** Thai license plate; Character segmentation; Character recognition; Pattern matching; Video stream

## Introduction

A Thai license plate has a rectangular shape which consists of 2 lines, the upper line and lower line. The upper line is divided into 2 parts, the first of which consists of 2 characters that can be a character and a number, and the second of which consists of 4 numbers. The upper line gives identification of the car. The lower part shows the name of province in Thailand in which the car license has been registered.

The plate is 15 by 34 centimeters in size, with a colored and embossed outline. The registration ID consists of two series letters followed by a serial number string of up to four digits, from 1 to 9999, without leading zeros, e.g. "กข 1" or "กข 1234". A leading number may be added in front of the two letters if the letter pool has been exhausted, as is the case in Bangkok since 2012, giving the format "1กข 1234". Both license plate styles are shown in Figure 1(a) and 1(b). Due to the case, the new plates, since 2012, have reduced text size to keep the plate size smaller. The province of registration is displayed in Thai underneath the registration ID. The colors of the text and background depend on the type of a vehicle.

The main problems of this project were how to find a place to record the video clip of many cars because of the strong security issue for an important building or place where there are many cars accessing the place, and how to reduce the noise of the image before processing. In Thailand many researchers were interested in license plate

recognition because of the serious problem of traffic jam in Bangkok but only a few of them were interested in license plate recognition from a video frame because the image from a video file lacks high precision. Due to the complexity of the operation method, we were encouraged to find a high-accuracy system of recognition of Thai license plate under an uncontrolled environment.


**Figure 1** Example of Thai license plate images (a) leading with 2 letters and 4 serial numbers style (b) leading with a number followed by 2 letters and 4 serial numbers style

There were three main processes in license plate character recognition, which were: 1) license plate localization, 2) character segmentation, and 3) character recognition (Babu and Reddy, 2013). Many researchers conducted car license plate recognition by using many methods.

**Artificial neural networks (ANN)**

This approach feeds character features in a car license plate into neural network nodes to identify a character. The ANN emulates knowledge of how the biological neural system stores and manipulates information. The notion is that artificial neural networks can solve all problems in automatic reasoning, including a pattern recognition problem. This approach classifies patterns by predictable properties of neural networks (Duangphasug and Thammano, 2006, Leelasantitham and Kiattisin, 2010, Mai et al., 2013, Shan, 2010, Nejati et al., 2013. Badr et al., 2011, Akoum et al., 2009 and Li and Guangli, 2011).

**Genetic algorithm (GA)**

This approach uses a probability or statistical pattern of character bit images in car license plate to generate GA chromosomes, which is used for classify a character. The input features of each character are extracted from a set of pattern characteristic measurements and transformed to a GA chromosome. Normally, a GA chromosome is represented by alphanumeric bit string, which length is depended on number of representing character features(Kim et al., 1996, Yin et al., 2008, Olivares et al., 2010 ).

**Template matching**

This approach uses a pattern of character bit images in car license plate to map with all characters templates in system database. The character templates are generated by extracting all characters image in car license plate training dataset. Ozbay and Ercelebi(2005) presented license plate identification by using a template matching method. The experiment was conducted on 340 car license plates, with the precision rate of 98.8 percent.

Dandu and Chopra (2012) proposed Indian vehicle number plates recognition by using the template matching and fuzzy logic system. The precision rate of the system was 95.1 percent.

Sa-ngamuang et al. (2007) showed Thai car license plate recognition by using an essential-element-based matching method. The experiment was conducted on 300 license plates, which consisted of 1719 characters, with the precision rate of 88.24 percent.

Based on related works, there were many researchers using many methods to recognize

characters of various languages from a vehicle license plate on a still image. This research tried to recognize a license plate under an uncontrolled environment on a video image. The details of the recognition method are presented in next sections.

## Materials and Methods

The experiment was conducted using the following computer hardware specifications: 1) CPU Intel® Core 2 Duo processors SU7300 (1.3 GHz. 800 MHz), 2) Memory DDR3 2 GB, and 3) Hard disk 100 GB. The computer software used 1) Microsoft Windows 7 (Microsoft Corp.; Redmond, WA, USA) as the operating system, 2) JRE (Java Runtime Environment 1.4 or higher), and 3) JMF (Java Media Framework 2.1.1e or higher) which is a Java library code that is suitable for using with video file added to Java application (Oracle Corporation; Parkway Redwood Shores, CA, USA). The digital video camera used was a Panasonic p2 HD (Panasonic Corporation; Oaza Kadoma, Kadoma-shi, Osaka, Japan). The process of analysis and design describes the system conceptual diagram and system structure chart. The details of each element are described below.

### System architecture

The TLPRS was a system of imitating reading abilities of a human. The system started with capturing cars license plates in a video frame and then used a computer system to recognize the characters in each license plate. Finally, the system displayed recognition results on a user's screen. The TLPRS system architecture is illustrated in Figure 2 (a).

### System structure chart

For more understanding, the TLPRS system architecture can be transformed into the system structure as shown in Figure 2 (b). The TLPRS consisted of 5 main modules, which were 1) video image acquisition, 2) plate localization 3) character segmentation, 4) character recognition, and 5) result presentation. The details of each module are described as the following:

***Video image acquisition***

The TLPRS used a digital camera to take video files of cars, which were running on a road, so the system could not control the environment, background, light, the distance between the camera and a car, and speed of each car, as shown in Figure 3 (a). The TLPRS captured a still image from a video file in every second. The video file consisted of 24 frames per second and 640 × 424 pixels per video frame.

***License plate localization***

This module consisted of 2 components, which were red-green-blue (RGB) to grayscale conversion, and black-and-white (B&W) conversion. The RGB to grayscale conversion component converted a still video image in the previous module. The NTSC standard RGB-to-grayscale image conversion was applied using Equation (1):

$$GR = 0.2989*R + 0.5870*G + 0.1140*B \quad (1)$$

where GR = grayscale image, R = red color, G = green color and B = blue color. The grayscale to B&W conversion was applied using a binarization technique. In this research, the grayscale image had value between 0-255 and the threshold value was 128, which means a grayscale value greater than 128 was considered as white. A sample of a B&W image is shown in Figure 3 (b). Then, the system searched for all possible license plates by marking rectangles over them, as shown in Figure 3 (c). Finally, the system searched for the proper aspect ratio (as shown in Equation 2) of all rectangles to detect a car license plate.

$$W = 1.2 * H \quad (2)$$

where W=plate width, H =plate height.

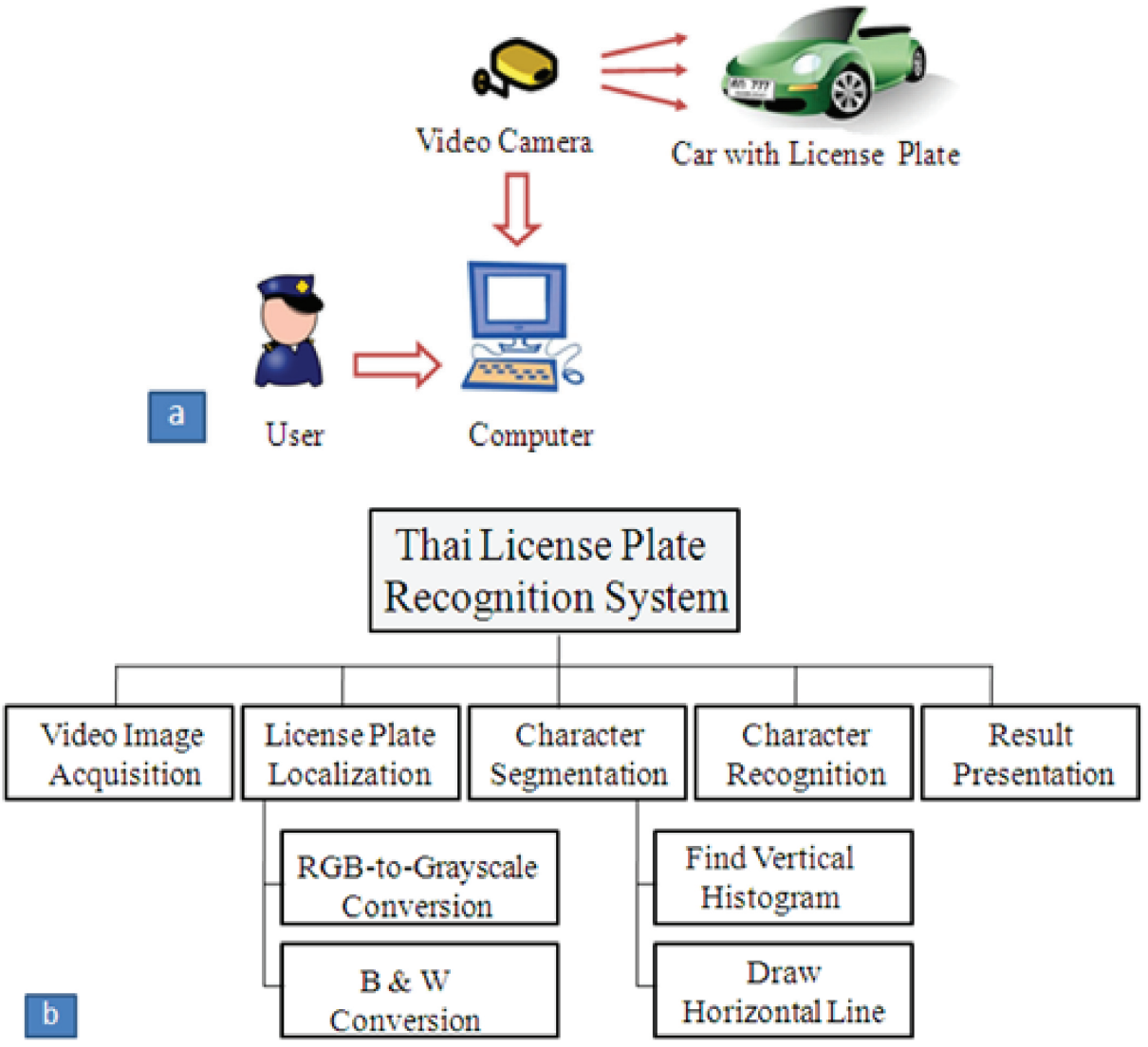


**Figure 2** (a) The TLPRS system architecture (b) the TLPRS structure chart


**Figure 3** The TLPRS sample images (a) RGB video image, (b) B&W image, (c) all possible license plate rectangles.

***Character Segmentation***

There were 2 components in this module, which were vertical histogram to separate each character, and drawing horizontal line to cross the middle of license plate characters. The sample license plate after localization is shown in Figure 4 (a) for the beginning of this module, and then the system applied a vertical histogram technique to separate each character in a license plate, as shown in Figure 4 (b). Finally, the system drew a horizontal line across the middle of characters in a license plate by counting alternative of B &W color more than 10 times. This research did not segment second horizontal line characters (province name) because of the characters were too small and not clear for segmentation and recognizing processes.

***Character Recognition***

The TLPRS applied a template matching technique to recognize characters on a license plate. Some samples of Thai characters and Arabic numbers on a training data set are shown in Figure 5 and number of training data set in each character is shown in Table 1. For a high recognition precision reason, the system resized both unknown characters and training characters to the same size first, and then compared black pixels of both characters.

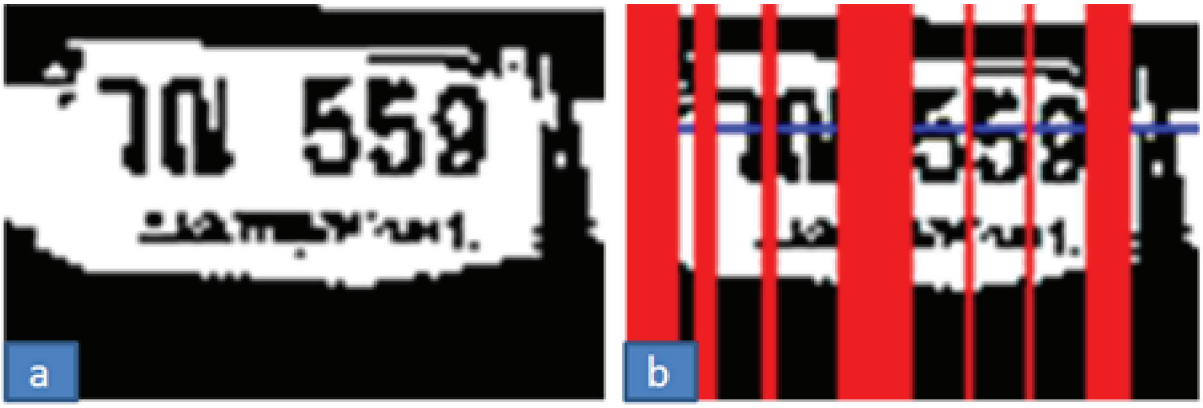


**Figure 4** The TLPRS segmentation sample (a) B&W license plate image (b) character segmentation by vertical histogram, including horizontal middle character line.

**Figure 5** Some Thai characters and Arabic numbers in Training data set

**Table 1** Number of Thai character and Arabic numbers in Training data set

| Character | Training number | Character | Training number | Character | Training number |
|---|---|---|---|---|---|
| ก | 4 | ธ | 3 | ส | 4 |
| ข | 2 | น | 3 | ห | 2 |
| ค | 3 | บ | 4 | พ | 1 |
| ง | 1 | ป | 1 | อ | 2 |
| จ | 1 | ผ | 1 | ฮ | 4 |
| ฉ | 2 | ฝ | 1 | 0 | 6 |
| ช | 4 | พ | 3 | 1 | 2 |
| ฌ | 2 | ฟ | 2 | 2 | 7 |
| ญ | 3 | ภ | 2 | 3 | 9 |
| ฎ | 3 | ม | 1 | 4 | 6 |
| ฐ | 2 | ย | 3 | 5 | 8 |
| ฒ | 2 | ร | 3 | 6 | 8 |
| ณ | 2 | ล | 1 | 7 | 5 |
| ด | 1 | ว | 5 | 8 | 8 |
| ถ | 3 | ศ | 9 | 9 | 5 |
| ท | 9 | ษ | 6 | Total | 169 |


**Figure 6** The TLPRS GUI screen (1) label 1 is video display box and (2) label number 2 is recognition result text box

***Result Presentation***

The graphic user interface (GUI) of TLPRS is shown in Figure 6. The system started with playing a video image file on the image box of GUI windows (Figure 6, label number 1), then the system automatically localized a car license plate in every second from a video file. After the system had localized the possible car license plate, the character segmentation routine separated each character. Finally, the recognition routine identified each character and the recognition results were displayed on text box below the image box playing a video file (Figure 6, label number 2).

**Results and Discussions**

The experimental was conducted at the Vibhavadee road intersection in Bangkok, Thailand during the daytime with normal light situation. The system took a video file length of about 2.5 hours, which consisted of 497 car images. The training dataset consists consisted of 169 alphanumeric characters, as shown in Table 1. The system precision rate is shown in Table 2. The system could localize 457 out of 497 car license plates and it could correctly segment 279 out of 457 localized license plates. The correct localization and segmentation of 279 license plates consisted of 1,646 alphabets, which were 531 Thai characters and 1,115 Arabic numbers. The system could recognize 477 out of 531 Thai characters and 1,058 out of 1,115 Arabic numbers. The character recognition precision rate of this system was 93.26 percent. The details of each character recognition rate are illustrated in Table 3.

There were many researchers conducted the Thai license plate recognition (Duangphasug and Thammano, 2006**,** Leelasantitham and Kiattisin, 2010 and Sa-ngamuang et al., 2007) but all of the previous researchers did not test with a video stream image under an uncontrolled environment. Therefore, this research tried to conduct the experiment with the applicable system by using a video frame format from the real roadside environment.

Some researchers conducted the license plate recognition system with a video stream (Nejati et al., 2013 and Dandu and Chopra 2012) but none of them conducted the experiment like our research, which used automatic localization, segmentation and recognition in one system.

There were 3 main processes in this research, which were license plate detection, character segmentation, and character recognition. Each process had the following difficulties. The location of license plate surrounded with many objects, special license plate back ground, and more than one vehicles in an image were three main problems in license plate detection procedure (as shown in Figure 7 (a), (b), and (c), respectively). A frame of license plate and dirty license plate were the two main problems for character segmentation procedure (as shown in Figure 7 (d) and (e), respectively). An unclear character was the main problem for character recognition procedure (as shown in Figure 7 (f) and (g)). An unclear character caused by 1) the speed of a running car, 2) a long usage of license plate and 3) some screws on license plate. Some successful license plate localization, segmentation and character recognition are shown in Figure 8 (a) - (g).

**Table 2** The license plate localization, segmentation and recognition precision of the system.

| Localization | | Segmentation | | Recognition | |
|---|---|---|---|---|---|
| **Correct** | **Incorrect** | **Correct** | **Incorrect** | **Correct** | **Incorrect** |
| 457 | 40 | 279 | 178 | 230 | 49 |
| 91.95% | 8.05% | 61.05% | 38.95% | 82.44% | 1.56% |

**Table 3** The recognition rate of each character.

| Character | Correct | Incorrect | Total | Character | Correct | Incorrect | Total |
|---|---|---|---|---|---|---|---|
| 0 | 84 | 6 | 90 | ฒ | 6 | 1 | 7 |
| 1 | 118 | 10 | 128 | ณ | 7 | 0 | 7 |
| 2 | 115 | 6 | 121 | ด | 0 | 0 | 0 |
| 3 | 130 | 0 | 130 | ต | 9 | 4 | 13 |
| 4 | 116 | 3 | 119 | ถ | 8 | 1 | 9 |
| 5 | 105 | 10 | 115 | ท | 27 | 6 | 33 |
| 6 | 97 | 5 | 102 | ธ | 4 | 2 | 5 |
| 7 | 86 | 4 | 90 | น | 7 | 0 | 7 |
| 8 | 106 | 9 | 115 | บ | 10 | 1 | 11 |
| 9 | 101 | 4 | 105 | ป | 0 | 3 | 3 |
| ก | 18 | 1 | 19 | ผ | 1 | 2 | 3 |
| ข | 4 | 0 | 4 | ฝ | 1 | 0 | 1 |
| ฃ | 0 | 0 | 0 | พ | 16 | 1 | 17 |
| ค | 12 | 1 | 13 | ฟ | 2 | 0 | 2 |
| ฅ | 0 | 0 | 0 | ภ | 7 | 1 | 8 |
| ฆ | 0 | 0 | 0 | ม | 9 | 1 | 10 |
| ง | 9 | 1 | 10 | ย | 12 | 2 | 14 |
| จ | 13 | 1 | 14 | ร | 12 | 1 | 13 |
| ฉ | 9 | 0 | 9 | ล | 12 | 1 | 13 |
| ช | 37 | 2 | 39 | ว | 36 | 2 | 38 |
| ซ | 0 | 0 | 0 | ศ | 44 | 4 | 48 |
| ฌ | 26 | 5 | 31 | ษ | 21 | 2 | 23 |
| ญ | 16 | 4 | 20 | ส | 19 | 2 | 21 |
| ฎ | 23 | 0 | 23 | ห | 7 | 0 | 7 |
| ฏ | 0 | 0 | 0 | ฬ | 5 | 0 | 5 |
| ฐ | 8 | 1 | 9 | อ | 6 | 1 | 7 |
| ฑ | 0 | 0 | 0 | ฮ | 14 | 0 | 14 |


**Figure 7** The TLPRS some incorrect process images, (a) – (c) incorrect license plate detection, (d) and (e) some incorrect character segmentation, and (f) and (g) some incorrect character recognition.


**Figure 8** The TLPRS some successful license plate localization, segmentation and character recognition images

**Conclusion**

This research fulfilled its objective which was to develop a computer system that had abilities of automatic localization, segmentation and recognition of characters on a vehicle license plate in a video stream. Although there were many researches on license plate recognition system, there were just a few license plate recognition systems of Thai characters under an uncontrolled environment. So this research focused on automatic license plate recognition on a video stream. The experiment was conducted on video files length of 2.5 hours, which consisted of 497 car images. The precision rate of the system was 93.26 percent.

However, there were many problems in doing this research, for example, the speed of running cars, the number of cars on a video frame, many types of vehicles on a road, inconsistent car license plate location, dirty license plates etc., which made it more difficult to recognize characters on a car license plate.

**References**

Akoum, A., Daya, B., and Chauvet, P., (2010) Two neural networks for license number plates recognition. *Journal of Theoretical and Applied Information Technology*. 12(1/2): 25-32.

Babu, G. S. and Reddy, K. C. S. (2013) Automatic license plate recognition system. *Journal of Engineering Research and Applications* 3: 870-876.

Badr, A., Abdelwahab, M. M., Thabet, A. M., and Abdelsadek, A. M. (2011) Automatic number plate recognition system *Annals of the University of Graiova, Mathematics and Computer Science Series*. 38(1): 62-71.

Dandu, B. R. and Chopra, A. (2012) Vehicular number plate recognition using edge detection and characteristic analysis of national number plates. *Journal of Computational Engineering Research*. 2: 795-799.

Duangphasug, P. and Thammano, A. (2006) Thai vehicle license plate recognition using the hierarchical cross-correlation ARTMAP. In *Proceeding of International IEEE Conference of Intelligent System*, 652-655, London.

Kim, S. K., Kim, D. W., and Kim, H. J., (1996) A recognition of vehicle plate using a genetic algorithm-based segmentation. In *Proceeding of International Conference on Image Processing,* 661-664, Lausanne.

Leelasantitham, A. and Kiattisin, S. (2010) A position-varied plate utilized for a Thai license plate recognition. In *Proceeding of the society of instrument and control engineers annual conference*. 3303-3307, Taipei, Taiwan.

Li, L. and Guangli, F. (2011) The license plate recognition system based on Fuzzy theory and BP neural network. In *Proceeding of International Conference on Intelligent Computation Technology and Automation*. 267-271, Guangdong, China.

Mai, V. D., Miao, D. and Wang R. (2013) Vietnam license plate recognition system based on edge detection and neural networks. *Journal of Information and Computing Science*, 8(1): 27-40.

Nejati, M., Pourghassem, H. and Majidi, A. (2013) Iranian license plate character recognition using mixture of MLP experts. In *Proceeding of International Conference on Communication System and Network Technologies*. 219-223, Gwalior, India.

Olivares, J., Palomares, J. M., Soto, J. M., and Gamez, J. C. (2010) License plate detection based on genetic neural networks,

morphological, and active contours. *LNAI* 6098: 301-310.

Ozbay, S. and Ercelebi E. (2005) Automatic vehicle identification by plate recognition. *World Academy of Science, Engineering and Technology*. 9: 222-225.

Sa-ngamuang, P., Thamnittasana, C., and Kondo, T., (2007) Thai car license plate recognition using essential-elements-based method. In *Proceeding of Asia-Pacific Conference on Communications,* 41-44, Bangkok, Thailand.

Shan, B. (2010) License plate character segmentation and recognition based on RBFneural network. In *Proceeding of International Workshop on Education Technology and Computer Science*. 86-89, Wuhan, China.

Yin, Z. J., Rui, Z. R., Min, L., and Yin, L. (2008) License plate recognition based on genetic algorithm. In *Proceeding of International Conference on Computer Science and Software Engineering*. 963-968, Hubei, China.