NUMBER PLATE RECOGNITION OF MALAYSIA VEHICLES USING SMEARING ALGORITHM

[1] .Munaisyah Abdullah, [2] Muayad Ali Hamood Bakhtan, [3] Shamsul Anuar Mokhtar
[1,2,3] Malaysian Institute of Information Technology
[1,2,3]University Kuala Lumpur, 1016, Jalan Sultan Ismail, 50250 Kuala Lumpur
For Correspondence; [1] munaisyah@unikl.edu.my, [2] muayad.ali@s.unikl.edu.my, [3] shamsulanuar@.unikl.edu.my

ABSTRACT: *Number Plate Recognition (NPR) is the extraction of vehicle license plate information from an image. The NPR uses the camera to take images. The quality of the images is a major factor in the success of the NPR. It should be generalized to process both of Malaysian number plates, single row and double rows. These plates usually written in English, and use different fonts, all plates may have a black background and white alphanumeric. In this paper, we present an efficient NPR system based on image processing using a Smearing algorithm for detection and segmentation. Template matching technique is used for the character recognition. The recognition outputs compared with the records in a database to come up with the specific information such as the vehicle's owner, place of registration, address. The system is implemented and simulated in Matlab. The system performance has been tested by 150 images of Malaysian plates based on the on the factors targeted in this research. It is observed from the experiment that the developed system successfully detects and recognize the vehicle number plate on real images.*

Keywords: Number Plate Recognition; Image Processing; Plate Recognition; Malaysian Plate; Smearing Algorithm, Segmentation, Template matching.

1. INTRODUCTION

Number Plate Recognition (NPR) is an essential subject of research due to its various fields of applications. Recognition of vehicle's number plate as reported in the literature, in general composed of several processing steps: extraction of the plate region, segmentation of characters, and recognition of each character [1]. The increasing numbers of vehicles draw a positive implication to provide an efficient NPR system. Hence, number plate recognition is needed to support the NPR applications [2]. As each and every vehicle will carry a unique number plate, so it is used as a personal identification for the vehicles [3].

NPR system has been used in various applications such as safe security system, parking access, private area identification, etc. The aim of this paper is to propose an efficient NPR system for Malaysian license plate for both single row and double rows. The proposed system is based on image processing using a smearing algorithm for detection and segmentation while template matching technique is used for character recognition. The template matching method has been used in the recognition process because it is sensible to find the difference between the characters.

2. RELATED STUDIES

There are various algorithms have been proposed to detect number plates in a different country. Besides that, plate detection has been studied a long time ago, but still a challenging task to detect it easily.

In Turkey, Smearing algorithm has been proposed on edge detection and some morphological operations such as erosion, filtering, and convolution. They applied the algorithm to scan the whole images in horizontal and vertical lines. The research achieved 92% accuracy [3], however, the system could not recognize few alphabets such as Q, X, and Y because these alphabets were not used in Turkish language alphabets.

Egyptian license plate recognition system has been presented. Two algorithms involved in this experiment which were Histogram Equalization with Sobel edge detector for detecting and Artificial Neural Networks (ANNs) for recognition. This system was tested by with 30 sample images and achieved 96% recognition rate [4].

In Malaysia, many studies have been done to recognize Malaysian license plate. In[5, 6], two new algorithms have been presented to detect and recognize Malaysian license plate number. The first algorithm called Speed up robust features algorithm (SURF), the second algorithm called Naïve Bayes algorithm. Both algorithms successfully detect and recognize Malaysian plate and the results reach 90-98% accuracy in plate detection, but in plate recognition it only achieved 67% accuracy. The main reasons that affect the system performances were due to the capturing noises and the environment.

Edge detection has been proposed with other algorithms such as Unwanted Line Elimination Algorithm ULEA, Chin code, Radon transform, Multilayer Perceptron MLP, Template matching and Machine vision[7-11]. Edge detection with ULEA applied in Malaysian license plate using 664 sample images reported 91%. Also applied in Iran license plate using 100, 1200 images and reported 98% and 99% accuracy respectively. As well as, its applied in Vietnam with neural network and MLP using Vietnam license plate and tested by 700 images and reported 97% accuracy. Ultimately, it's applied with template matching and machine vision using 130 images of china license plate and reported 78% accuracy. Susceptibility of edge detection algorithm to integrate with other algorithms, make them widely used in many countries.

Template matching with Support Vector Machine with neural network were used successfully recognize Indian license plate with 96% recognition rate[12]. Advantages of this study in the template matching method used digits 0-9 and alphabets a/z & A/Z respectively. A similar database is needed to compare all edges-based methods.

Edge detection and Soble edge detection have been proposed to recognize non-standard Malaysian car license plate for both single row and double rows. The system was tested

using 500 sample images and resulted in 64% recognition rate using Malaysian plate single row and 51% recognition rate using Malaysian plate double rows. The main reason for the poor performance of the algorithms was due to the big variation between the characters, plate sizes and font types[13].

Edge detection with data extraction has been proposed for recognizing Pakistani license plate to be used for security purposes[14]. This experiment tested using 125 different images and reported 81% accuracy.

License plate recognition system based on pulse coupled neural networks and template matching has been studied in China[15]. Additionally, a morphological operation was also involved in this experiment especially in dilation and erosion. The method has some advantages when the image is polluted or photographed without enough illumination, the system is able to detect it. However, the limitation of this method is the incapability of the system to convert text or data accurately, thus it needs to be improved. Besides, the characters, background and size of the vehicle plate vary from one country to another.

Three algorithms including Edge detection, Scale-Invariant Feature Transform (SIFT) and template matching has also been proposed to recognize Malaysian license plate. The research was done using 500 Malaysia plate images and reported 81% recognition rate[16]. The candidates were in a good condition. Nevertheless, blurring and low-resolution images could not be dealt with successfully.

Guessing filtering with edge detection and morphological operations has been proposed and the experiment was tested using 250 Australian license plates and achieved 96% accuracy[17]. The proposed technique is effective in real-world applications due to the regularity of Australian license plate.

From above literature, it can be concluded that license plate recognition still an open research topic. However, many studies have been done on plate recognition system, but still need some improvements. On the other hand, many algorithms have been proposed, while in this experiment smearing algorithm is the most suitable for detecting Malaysian plates.

3. METHOD

NPR PROCESS

The NPR system consists of four major processes which are preprocessing, plate localization, character segmentation and character recognition. The flow of the processes is illustrated in Figure 1.

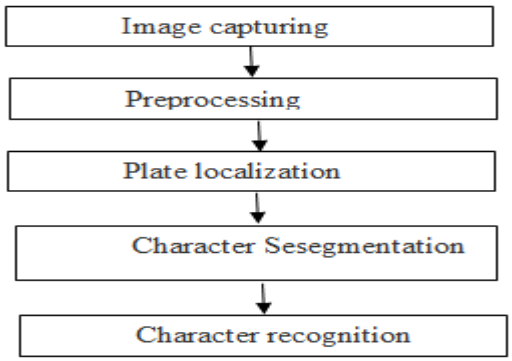

Fig 1 : Number plate Recognition System processes

First of all, the image is captured in RGB format and the resolution is 800 x 600 pixels. In this research, Malaysian plates (single row and double rows) were the sample data, as shown in figure 2 tow image capturing for Malaysian plates.

Fig2. Image Cappturing by USB Camera

In the second process, the image will be enhanced by converting it to the binary; the main objective of this process is to reduce the number of the pixels in order to ease the future processes. The binary process shown in figure 3.

Fig3. Binary Image

After that, the exact location of the plate is extracted by scanning the plate region using a horizontal and vertical algorithm (smearing algorithm). After the localization of the vehicle plate comes out, the next process will be applied which is character segmentation. It can be observed that the smearing algorithm has successfully detected the Region of Interest (ROI) that only contains vehicle number plate as shown in figure 4.

Fig 4. Plate Region Extracting Using Smearing

Character segmentation is an important step. It is a process that separates words to a single character for easy identification[18]. There are numerous factors that can affect the system performance in this step, such as the influence of image's noise, the frame of the license plate, rivet, space mark, vertical overlap, horizontal overlap ligature and etc. In this process, the row and column segmentation methods were used to extract the individual character in the vehicle number plate as shown in figure 5.

Fig 5. Character Segmentation Process

Character recognition depends on template matching. It is one of the widely used tools for character recognition. This technique has the advantage for recognition of the single font, not rotated, not tilted, and fixed size characters. Although this approach preferably utilized with binary images. Template matching method must have character images as template stored in memory. The segmented character will be identified by calculating the correlation coefficient. The idea behind the implementation of a correlation based identification scheme is when two template pools, one consisting of all the possible values of letters, and of all values of the digits, are constructed.

In this research, template matching and cross-correlation function was used, each alphanumeric character can be recognized as shown in figure 6.

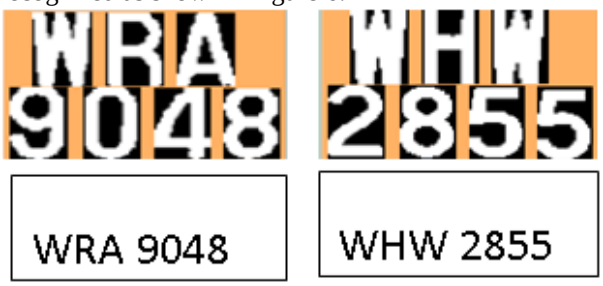

Fig 6. Character Recognition Process

Data Classified

To test this experiment, 150 sample images of Malaysian plates are included. The images are taken using Samsong Camera in different condition based on the factors targeted in this research which are standard or non-standard Malaysian plates, clarity of the image or the noises, Illumination (good lighting- poor lighting), distance, and tilt angle.

According to the Road Transport Department (Jabatan Pengangkutan Jalan (JPJ)), the standard group implies that the font style of the plate number is italic while the non-standard group has not italic font style

In addition, other factors such as different effects of the day lights and environments were also considered in this research. The sample images were taken from a distance (D) ranging between 0m and 3.5m. These sample images can be classified into four groups based on the distance (D).

Besides that, the images can also be classified into two groups based on tilt angle (Ø). The first group includes all images that were taken with a tilt angle (Ø) between 0 and 20 degrees, while the second group includes all images between 20 and 40 degrees.

4. RESULTS AND DISCUSSION

There are 150 samples of images that have been captured using a Samsung camera with a different resolution, lighting, and environments. The success rate of each factor will be discussed in this section. The equation has been used to calculate the recognition errors as presented in [19].

$$E= (1-(\sum 1-T) / C) \quad (1)$$

Whereas T is the total of wrong character recognition and C is constant = 7 which represents total alphabets in the plate. The recognition accuracy is calculated using this equation :

$$R= (1-E) * 100 \quad (2)$$

Where R is the recognition accuracy, E is the total Errors. Table 1 shows the results and the accuracy in each process in this experiment.

Table 1: The Success Rate Of The Npr System In Each Process

The process	Success rate%	Error %
Plate localization	97.4	2.6
Character Segmentation	96.0	4
Character Recognition	76.0	24

As shown in table 1 the experiment results have achieved an overall plate localization rate of 97.4% , while the success rate of the segmentation process has achieved 96% . We achieved an overall plate recognition rate of 76.0% .

From the testing results, we observed that the character recognition process has a lower present age. In addition, characters and numbers affected by the rivers of plates have gotten a higher percentage of misclassification. Table2 shows all factors that affected the system performance.

Table 2: The Affect Factors That Targeted In This Research

Factors	Plate Localization%	Character Segmentation%	Character Recognition	Success Rate
Standard	99.0	88.0	82.0	89.6
Non-Standard	88.0	96.0	70.0	84.6
Clear	99.0	98.0	79.0	92.0
Noise	97.0	95.0	74.0	88.6
Good illumination	99.0	96.0	76.0	90.3
Poor illumination	83.0	95.0	75.0	84.3
Distance $0< D\leq 1$	95.0	94.0	80.0	89.6
Distance $1< D\leq 2$	99.0	98.0	76.0	91.0
Distance $2< D\leq 3$	98.0	97.0	69.0	88.0
Distance $3< D\leq 3.5$	94.0	95.0	65.0	84.6
Angle $0\leq Ø\leq 20$	97.0	96.0	76.0	89.6
Angle $20<Ø\leq 45$	94.0	89.0	74.0	85.6

It can be seen from the table above the success rate indicated that 89.6% is appropriately detected and recognized by the system for standard plate and 84% for non-standard plate. This variation can be minimized if we add other futures like car shape or other references.

On the other hand, the noise of the plate has affected the performance of the system as shown above in Table 2. We can see the variation appears slightly in the segmentation process which reaches 95% accuracy comparing with segmentation in the clear image which achieved 98%.

In the recognition process, the variation is significantly higher, because the system cannot differentiate between the character's edges and the spots in the plate.

The experiment results have shown that the accuracy for the factor of Good illumination appears to be higher against the factor of Poor illumination in all processes. It can be seen from Table 2, the success rate achieved 90.3% is appropriately detected and recognized by the system for the good illumination and 84.3% for the poor illumination. This variation can be decreased by controlling the lighting or use the special camera that has high resolution.

The performance of the system is tested against the distance (D) between the camera and the car plate number. The experimental results have shown that the highest success rate 91.0% is achieved within the second group (1<=D<2), followed by the first group (0<D<1) with a success rate of 89.6%. However, it is also shown that the larger the distance (D) between the camera and the car plate, the lower the success rate of the system, as shown in the third and fourth groups. The findings of this factor prove that the proposed system has a good success rate ranging between 84.6% and 91.0% with a distance between the camera and the car plate ranging between 0 meter and 3.5 meters, while the highest success rate is achieved when the distanced is between 1 to 2 meters.

The performance of the system is also tested against the tilt angle (Ø). The results listed in Table 2 show that the success rate of the first group (0<Ø<=20) is 89.6%, while the second group (20<Ø<=40) produced a lower success rate of 85.6 %. It was clear that the performances of localization and segmentation process are relatively good compared to the performance of the recognition process. The segmentation process is affected by the tilting of the characters since characters' boundary boxes are rectangular in shape, while the tilted characters are generally not affected. The recognition process, however, has the lowest performance of 76.0% and 74.0% in the first and second groups respectively. This is due to the fact that tilted plate numbers have tilted characters which are not geometrically similar when compared and matched with the reference characters in the recognition process.

The results shown that the developed NPR algorithm has successfully detect the number plates in various day conditions. It can also detect and recognize vehicle plates with various factors that targeted in this research, such as standard and non-standard number plate, distances, angle, etc. For example, distance affects the size of the number plate in an image. Once the vehicle number plate is detected, the individual characters are recognized using template matching and cross-correlation function. The correlation method has a high efficiency of the character recognition process and the probability of the recognition can be calculated.

5. CONCLUSION

In this paper, the vehicle identification system is presented using their number plates. The system uses a series of image processing techniques for identifying the vehicle license plate number. The system is implemented in Matlab and its performance is tested on real images. The experiment results show that the system robustly detect and recognize the vehicle using number plate against different conditions and analyses the main factors that affect the system performance. Nevertheless, there is still room for improvement; the camera used in this experiment is a casual camera. The system robustness can be increased if high-resolution camera is used. The system is computationally inexpensive and can be implemented for real time vehicle identification system in the future.

6. REFERENCES

[1] X. Jia, X. Wang, W. Li, and H. Wang, "A novel algorithm for character segmentation of degraded license plate based on prior knowledge," in *2007 IEEE International Conference on Automation and Logistics*, pp. 249-253 (2007).

[2] H. A.-Q. Maarif and S. Sardy, "Plate number recognition by using artificial neural network," *Prosiding Semiloka Teknologi Simulasi dan Komputasi setra Aplikasi (* 2006).

[3] S. Ozbay and E. Ercelebi, "Automatic vehicle identification by plate recognition," *World Academy of Science, Engineering and Technology*, vol. **9**, pp. 222-225, (2005).

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

[5] F. M. Khaleel, S. N. H. S. Abdullah, and M. K. B. Ismail, "License plate detection based on Speeded Up Robust Features and Bag of Words model," in *Smart Instrumentation, Measurement and Applications (ICSIMA), 2013 IEEE International Conference on*, pp. 1-5 (2013).

[6] A. A. Halin, N. M. Sharef, A. H. Jantan, and L. N. Abdullah, "License plate localization using a Naïve Bayes classifier," in *Signal and Image Processing Applications (ICSIPA), 2013 IEEE International Conference on*, pp. 20-24 (2013).

[7] A. M. Al-Ghaili, S. Mashohor, A. R. Ramli, and A. Ismail, "Vertical-edge-based car-license-plate detection method," *IEEE transactions on vehicular technology*, vol. **62**, pp. 26-38, (2013).

[8] R. Azad and M. Baghdadi, "Novel and Fast Algorithm for Extracting License Plate Location Based on Edge Analysis," *arXiv preprint arXiv:1407.6496, (*2014).

[9] R. Azad and M. Nazari, "Novel and Automatic Parking Inventory System Based on Pattern Recognition and Directional Chain Code," *arXiv preprint arXiv:1407.6321, (*2014).

[10] V. Mai, D. Miao, and R. Wang, "Vietnam License Plate Recognition Systembased on Edge Detection and Neural Networks," *Journal of Information and Computing Science*, vol. **7**, pp. 27-40, (2013).

[11] J. Yang, B. Hu, J. Yu, J. An, and G. Xiong, "A license plate recognition system based on machine vision," in *Service Operations and Logistics, and Informatics (SOLI), 2013 IEEE International Conference on*, pp. 259-263 (2013).

[12] K. Bhosale, J. Jadhav, S. Kalyankar, and R. Bhambare, "Number Plate Recognition System for Toll Collection," *International Journal of Emerging Technology and Advanced Engineering*, vol. **4**, pp. 729-732, (2014).

[13] H. Y. Chai, H. H. Woon, L. K. Meng, and Y. S. Li, "Non-standard Malaysian car license plate recognition," in *Computer Applications and Industrial Electronics (ISCAIE), 2014 IEEE Symposium on*, pp. 152-157 (2014).

[14] M. A. Saeed, M. Waqas, Z. Akbar, and N. Ali, "An Efficient way of Number Plate Alphabets and Numbers Extraction for Security Purpose."

[15] X.-h. Wang, J.-j. Yu, Z.-h. Miao, and Y. Song, "License plate recognition based on pulse coupled neural networks and template matching," in *Control Conference (CCC), 33rd Chinese*, 2014, pp. 5086-5090 (*2014*).

[16] H. Rajput, T. Som, and S. Kar, "An Automated Vehicle License Plate Recognition System," *Computer*, vol. 8, pp. 56-61, (2015).

[17] P. Pandya and M. Singh, "Morphology based approach to recognize number plates in India," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. **1**, pp. 107-112, (2011).

[18] B. Al-Badr and S. A. Mahmoud, "Survey and bibliography of Arabic optical text recognition," *Signal processing*, vol. **41**, pp. 49-77, (1995).