SEGMENTATION AND RECOGNITION OF MALAYSIAN CAR PLATES USING FREEMAN CHAIN CODES (FCC)

NOR AMIZAM BINTI JUSOH

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Master of Science in Computer

Faculty of Computer Systems & Software Engineering
UNIVERSITI MALAYSIA PAHANG

FEBRUARY 2011

ABSTRACT

Research on automatic car plate recognition has been widely and intensively conducted all over the world with the application of various types of segmentation and recognition techniques by researchers. Most of the recognition techniques applied are not focusing on shape-based recognition although through normal human vision, each character in the car plate has a unique and different shape with each other. The chosen techniques have contributed into many proposed suitable methodologies for car plate recognition research but with the same objectives; to gain high or increase the segmentation and recognition accuracy rate with less processing time. This research is conducted with the aim of identifying the suitable or appropriate segmentation technique which can be used to segment either the standard or non-standard specification car plates. Besides that, the objective is also to study whether the shape-based recognition technique is efficient and accurate enough to recognize Malaysian car plates which are varied in terms of font types. Techniques that have been chosen for segmentation process are pixel count, connected component labeling (CCL) and a proposed technique; the connected component labeling with minimum object removal. As for recognition purpose, the techniques that have been experimented are the Freeman chain codes (FCC), template matching and a proposed technique; the Freeman chain codes with characters' features (FCCwF). The results from the experiment shows that the proposed segmentation technique; the connected component labeling with minimum object removal able to increase the segmentation success rate by more than 96% and the proposed recognition technique; the Freeman chain codes with characters' features is able to reach the accuracy rate of 95% compared to other tested techniques. Based on the experiments and results, Freeman chain codes are efficient and accurate enough to recognize various font types of Malaysian car plates for most of the characters with less processing time of 0.1s but a higher recognition accuracy rate can be gained by combining FCC with characters' features.

ABSTRAK

Kajian terhadap segmentasi dan pengecaman nombor plet kereta secara automatik telah lama dijalankan diseluruh dunia dengan penggunaan pelbagai teknik dan kaedah oleh penyelidik-penyelidik. Namun begitu, kebanyakan pilihan teknik pengecaman tidak tertumpu kepada pengecaman berasaskan bentuk sedangkan secara amnya, setiap huruf yang terdapat di dalam plet kereta itu mempunyai bentuk yang unik dan berbeza antara satu sama lain. Pilihan teknik pengecaman ini menghasilkan variasi metodologi tetapi dengan maksud dan tujuan yang sama iaitu mencapai kadar ketepatan pengecaman yang tinggi dengan masa memproses yang kurang. Oleh itu, kajian ini telah dijalankan bagi mengenalpasti teknik segmentasi yang sesuai yang dapat mengasingkan huruf-huruf di dalam plet kereta yang dibina sama ada mengikut spesifikasi piawai ataupun tidak. Selain itu, kajian juga bertujuan untuk mengkaji sama ada teknik pengecaman berasaskan bentuk adalah sangat berkesan dan tepat dalam mengecam plet kereta Malaysia yang dihasilkan menggunakan pelbagai jenis bentuk tulisan. Teknik yang telah dipilih untuk proses segmentasi adalah teknik kiraan piksel (*pixel count*), pelabelan komponen berhubung (*connected component labeling*) dan teknik cadangan iaitu pelabelan komponen berhubung dengan penyingkiran objek bernilai minima. Manakala, teknik yang dipilih untuk eksperimen pengecaman adalah teknik kod rantaian Freeman (*Freeman chain codes*) yang ringkas yang menghasilkan kod rantaian berdasarkan piksel sempadan bagi imej huruf, padanan templat (*template matching*) dan teknik cadangan iaitu teknik kod rantaian Freeman dengan ciri-ciri huruf. Hasil yang diperolehi dari kajian menunjukkan bahawa teknik segmentasi cadangan iaitu pelabelan komponen berhubung dengan penyingkiran objek bernilai minima mampu meningkatkan tahap kejayaan segmentasi melebihi 96% dan teknik pengecaman cadangan iaitu teknik rantaian Freeman dengan ciri-ciri huruf dapat meningkatkan kadar ketepatan pengecaman sehingga 95% berbanding teknik-teknik ujian lain. Berdasarkan hasil tersebut maka, teknik kod rantaian Freeman boleh menjadi satu alternatif untuk mengecam plet kereta Malaysia dengan pelbagai jenis bentuk tulisan dengan kemampuan teknik ini untuk memberikan tempoh masa memproses yang kurang iaitu 0.1 saat dan kadar ketepatan pengecaman yang tinggi dapat dicapai melalui gabungan bersama ciri-ciri huruf.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CCL	Connected component labeling
CPR	Car plate recognition
FCC	Freeman chain code
FCCwF	Freeman chain code with features
N/A	Not available
RBF	Radial basis function
TM	Template matching

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

Digital image processing refers to processing digital images by using digital computer in order to change its nature such as enhancing the edges of an image to make it appear sharper, to remove noise from image, to remove detail from image and etc. Digital image processing needs a very appropriate and fast technique especially when dealing with images which involved in detection and recognition tasks. One of the major advantages in digital images is the ability to perform a variety of processing procedures with a computer. These procedures can be selected and adjusted to change the characteristics of the images, usually for the purpose of improving quality or optimizing characteristics for maximum visibility. Images can be stored in the computer memory and easily retrieved on the same computer screen and can be saved indefinitely or be printed on paper or film if necessary. All digital imaging systems can be networked into practice management software programs facilitating integration of data. Digital imaging is also environmentally friendly since it does not require chemical processing. It is well known that used film processing chemicals contaminate the water supply system with harmful metals such as the silver found in used fixer solution. Apart from all of the above advantages, the use of digital image processing also has disadvantages. One of the advantages is the initial cost can be high depending on the system used, the number of detectors purchased or the hardware or device used. Competency of using the software can takes time to master depending on the level of computer literacy of team members.

Many applications have been developed and implemented which applies image processing and character recognition technology. One of the most popular and important applications of digital image processing which focus on image detection and recognition is car plate recognition (CPR) system which is used to overcome the challenges of monitoring modern day traffic. This technology identify vehicles automatically by reading and recognizing their car plates which may involve in traffic violations, crime scenes, car parks area determination, statistical and etc. The plate recognition is done based on the given conditions and instructions. This recognition system is installed in many places such as toll gates, parking lots and also entrance of highly secured buildings. Police are using this application because they can detect speeding vehicles from distance away and summon tickets will be issued later to the vehicle's owner. These systems are beneficial because it can automate car park management, improve the security of car park operator and the users as well, eliminate the usage of swipe cards and parking tickets, improve traffic flow during peak hours, detect speeding cars on highways, and detect cars which run over red traffic lights. Furthermore, when the data gathered by a CPR system is stored and organized in a database, more complex information-driven tasks may be achieved, such as vehicle travel time calculations, marketing analysis, and border control.

Different standards for car plates are found all over the world, and variations within each country are usually limited. Because of that, many commercialized systems to identify car plates have been developed which are mainly from Japan, USA and Europe but the different styles and standards of car plates used in these countries caused these commercialized systems are unsuitable to be used in other countries especially in Malaysia. Automated car plate reading is a particularly useful and practical approach because, apart from the existing and legally required car plate, it assumes no additional means of vehicle identity.

The research and development (R&D) for car plate recognition system has been carried intensively by many researchers all over the world. It has started since 1989 and continues to evolve until today by the application of various techniques ranging from simple to complex technique. One of the simple techniques applied in CPR was Freeman chain code (FCC) which is done by Paras Ram (2005) but has not been

published in any journal or conference eventhough the recognition result achieved was quite impressive. Recent researches on CPR have been performed using neural networks by Shan (2010) which achieve high recognition rate for normal characters but not Chinese characters. So, this has become a motivation to continue research on CPR since FCC is a shape-representation technique which is very good in representing shapes or objects information with curves. The advantages of this technique also can be considered in proposing FCC as an alternative recognition technique.

Normally, when conducting R&D in this area, a general methodology which consists of several phases is applied together with the appropriate techniques in order to ensure that the R&D will meet its objectives. The general phases are image acquisition, image pre-processing, image segmentation and image recognition and each phase is performed continuously. Image acquisition is a phase for obtaining samples of images while image pre-processing is a phase for enhancing the quality of images obtained from the first phase. Image segmentation is a phase to segment images whether to isolate object from background or to separate characters on car plates. The final phase is image recognition where in this phase, the separated image is recognize and in this case, the characters on car plates.

1.1 RESEARCH BACKGROUND

Car plate recognition is important because of its advantages as already discussed in previous section. The choice of techniques for each phase especially the segmentation and recognition techniques plays an important role because it will ensure that the information extracted is correct. The information extracted is the characters that depict the car plate. The segmentation technique chosen is normally based on the capability to isolate objects from its background successfully while the recognition technique is based on the capability of the technique to give a high recognition rate, recognize similar pattern characters and less processing time which is the time used to process and recognize car plates. Certain recognition techniques show a high recognition rate while others show less processing time but the best recognition technique should deliver both advantages.

Image segmentation is not an easy task and it is quite difficult to perform because of several reasons such as non-uniform illumination, no control of the environment, inadequate model of the object of interest and noise in the image. Due to this, to ensure the successfulness of segmentation phase, the images feed into this phase must be in high quality. If the images are not in a good quality, it must go through certain processes in order to remove all noises which will affect the segmentation result. Segmentation affects greatly on the accuracy of recognition because if the contours of the characters are inaccurate, it might lead to errors in the recognition phase or even failed the recognition phase. One of the factors which can makes the contours of characters inaccurate is slope (Al Faqheri and Mashohor, 2009) which sometimes happened during image acquisition phase. This factor can be dealt by applying vertical and horizontal scanning (Al Faqheri and Mashohor, 2009). Since Malaysian car plates are varied in terms of its types, fonts, sizes and styles, therefore, to apply only one segmentation technique is not a suitable solution (Wu et al, 2004). This research investigates on the suitable technique to apply with the variation types of normal single row Malaysian car plates.

In normal human vision, characters in car plates are differentiated by observing its shape characteristics whether the outer or edge shape or with its region. The shape characteristics can be classified into three classes; contours such as chain codes technique, regions such as template matching technique and transforms such as wavelet technique (da Fontoura Costa and Cesar, 2009). These three classes differ in terms of how the shape analysis is done and represented. Eventhough the characters can obviously be differentiated by their shape, however, from the researcher's findings on recognition techniques for car plate recognition system (CPR), various techniques have been applied by researches besides the shape-based representation technique. The recognition techniques are varied from simple to complex techniques and some of the researchers are focusing on the application of template matching technique (Naito et al, 2000; Sarfraz et al, 2003; Tao et al, 2004; Wu et al, 2005; Wang and Zhang, 2008; Quan et al, 2009) while only one research has been conducted using stroke analysis or Freeman chain code (FCC) (Paras Ram, 2005). The evolution of researches on CPR with its recognition techniques applied can be referred in Section 2.3.

FCC is a technique which uses only a boundary of an object in a clockwise or anti-clockwise direction and assigning a direction to the segments connecting every pair of pixels by referring to either 4 or 8-neighbor. Since this technique only reads the boundary pixels and not the whole region of shape, so the time taken to process the information will be lower than the time used to read and process the whole regions. Therefore, with this benefit, FCC is assumed to be able to give high recognition rate and at the same time able to process or recognize car plates in high speed. Due to that, this research is conducted to study on the efficiency and accuracy of FCC as one of the alternative recognition technique. This technique is compared with the template matching technique in terms of recognition accuracy and processing time. If this technique alone does not yield high recognition rate, it will be combined with characters' features to improve the performance of recognition rate and at the same time preserve the advantage of less processing time.

1.2 PROBLEM STATEMENT

One of the major concerns for car plate recognition research is to apply a segmentation technique which is able to segment characters in car plates successfully and a recognition technique which is capable in giving high recognition accuracy and less processing time. For the recognition phase, both capabilities are assumed as able to be achieved by using the Freeman chain code technique; one of the shape-based representation technique as already discussed in general in previous section. According to Cha et al (1999), chain codes is one of the techniques that are able to recognize characters and digits successfully and any shape feature can be computed from the chain codes (Jahne, 2005) while Seul et al (1999) claims that chain codes provide a lossless compressing and preserving all topological and morphological information which bring out another benefit in terms of speed and effectiveness for the analysis of line patterns. From the findings, chain codes technique have been studied widely in other areas besides car plate recognition system such as handwritten word recognition (Madhvanath et al, 1999; Chan and Yeung, 1999; Britto et al, 2000), line drawing interpretation (Haron et al, 2004), object or image recognition (Salem et al, 2005), corner detection (Subri et al, 2006), length estimators (Dianat and Haron, 2008), recognizing handwritten

characters (Hoque et al, 2003; Shilbayeh et al, 2009) and eyes contour detection (Jing-Dong, 2010).

Thus, this research aim is to study on the efficiency and accuracy of Freeman chain code as an alternative recognition technique in car plate recognition. Specifically, this research is intended to answer the following questions:

a. What are the existing methodologies that have been used by other researchers in car plate recognition?
b. Are template matching and FCC techniques efficient or accurate enough for car plate recognition?
c. If not accurate, is there a way to make FCC more accurate?

1.3 OBJECTIVE

The objectives for this research are:

a. To experiment on the accuracy and effectiveness of FCC and compares it with template matching.
b. To develop a new algorithm to increase the recognition rate.

1.4 THE IMPORTANCE OF THE RESEARCH

This research is to study, analyze and test the FCC technique, algorithm and methodology needed for automatic recognition of car plates. It is hope that this research can contribute in the numbers of recognition techniques in image processing area. The experience, knowledge and advantage obtained from this research can help in image processing area especially in enhancing and improving the application of car plate recognition.

1.5 SCOPE

This research only focuses on:

a. Single row car plates with black background and white characters.
b. Normal type of Peninsular Malaysia's car plate with standard specification.
c. Images of car plates in JPEG format taken by using digital camera.
d. The image of car plate for research purpose.
e. Car plates with three alphabets and four numbers.
f. Characters to be recognized are ranged from A to Z with the exception of I,O and Z while the numbers will range from 0-9.

1.6 CONTRIBUTIONS

The academic contributions from this research are:

a. A methodology for car plate recognition which uses FCC technique and characters' features.
b. A segmentation technique and algorithm which combines connected component labeling (CCL) and minimum object removal.
c. A recognition technique and algorithm which combines FCC technique with characters' features.

1.7 THESIS ORGANIZATION

The organization of this thesis is consisting of six important chapters from Chapter 1 until Chapter 6. Chapter 1 presents the introduction of the research including the general review of previous approaches in the car plate recognition done by other researches, research background, issues related with the needs to conduct this research which is included in the problem statement section followed by the objectives, the importance of the research, scopes and the contributions of the research. Chapter 2 will explore in details about all related theoretical backgrounds, the evolution of research in CPR systems, previous works done by existing researches on CPR using FCC and

template matching. Chapter 3 will elaborate in detail how the research was conducted for the first module while the explanation on the experiment for module 2 is discussed in Chapter 4. The results and discussion is discussed in Chapter 5 and the conclusion and suggestions for the whole research as well as the future works that could be done to improve the system are discussed in Chapter 6.

1.8 SUMMARY

This chapter discusses on the emerging area of image processing and the R&D carried out under this area. One of the researches is car plate recognition system which has been intensively done all over the world. This application has many advantages such as helps the enforcer to enforce the law, helps in solving crimes cases and improve the security. In order to ensure the successfulness of the application, many techniques have been applied to increase the recognition rate, minimize the error rate and lower the processing time. Since characters are differentiated by its shape and not many researchers tend to use the shape-based representation technique for recognition, therefore the FCC technique has been chosen to be experimented to study on its efficiency and accuracy. The suitable segmentation techniques for Malaysian car plates also become the focus for this research because segmentation affects greatly the recognition rate.

CHAPTER 2

LITERATURE REVIEW

2.0 INTRODUCTION

Nowadays, recognition system plays an important role in many fields such as medical imaging, manufactured goods inspection, industrial inspection and law enforcement by applying image processing applications. One of the image processing applications is car plate recognition (CPR) system which developed using the image processing technique. This application has been widely used all over the world in assisting the law enforcement to recognize cars which involved in certain offences or crimes due to some benefits offered by this system (Casaletto, 2010). Normally this application is connected with cameras installed in patrol cars to track down auto-related crimes. These cameras can take pictures of car plates from distance and the pictures captured is processed using the image processing techniques and then run against a database of stolen vehicles to find possible matches. Besides that, CPR system also is widely used as a way to detect speeding vehicles on highways. Vehicles with speeds more than allowed speed can cause harm to the drivers himself or other people involved in accidents. In order to avoid such things happened, the CPR can be used to detect the speeding vehicles by capturing the images of vehicles' plates and processed it. The result from the processing can be used to print summon tickets to the speeding drivers.

Research on car plate recognition has been conducted intensively with the application of various techniques ranging from simple and single technique to a complex or combination of several techniques. The application of various techniques was due to several reasons or factors that support the emergence of the research in this area. The factors are such as to increase the recognition rate, minimize the error rate and

processing time. These three factors are very important to ensure the success of a CPR system to be considered as a reliable system. In this thesis, previous researches have been studied in order to gain some knowledge on the important aspects of car plate recognition research such as the methodology of the recognition system, the techniques and the algorithms applied. All theoretical parts of all elements needed for this research are discussed later in this chapter.

2.1 MALAYSIA'S CAR PLATE

Car plates in Malaysia are varied and different in colors and formats. The formats are single and double rows car plates and the color or background of car plates for personal and commercial vehicles are black with white characters on it. There are also car plates with red background and white or yellow characters which are provided for diplomats. Figure 2.1 shows the different types of Malaysia's car plates. Another type of car plates is the special or customized plates which have words such as 'BAM bee', 'SATRIA', 'PROTON', 'TIARA', 'PUTRAJAYA', 'SUKOM' and others as shown in Figure 2.2. For this research, only single row car plates for personal use as shown in Figure 2.3 are used as testing images. The Road and Transport Department of Malaysia has endorsed a standard for car plates that includes the font and size of characters that must be followed by car owners (Alias, 1999). Figure 2.4 and Table 2.1 show the standard specification for normal car plates while Appendix A shows standard specification for customized car plates.

Figure 2.1: Different types of Malaysian car plates

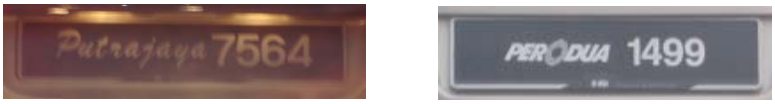

Figure 2.2: Special event or customized Malaysian car plates

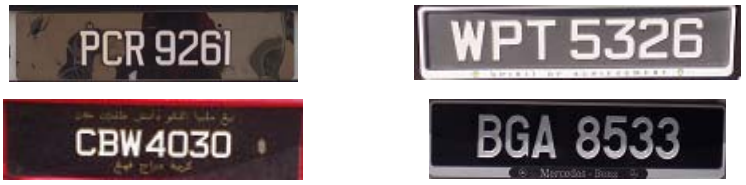

Figure 2.3: Different types and sizes of fonts used for normal Malaysian car plates.

Table 2.1: Specific plate measurement and criteria of car plate

No.	Specification	Measurement (mm)
1.	Character height	70
2.	Character width	40
3.	Distance between characters or numbers	10
4.	Distance between character and number in single row plate	30
5.	Distance between character row and number row in double row plate	10
6.	Minimum clearance between character and top-edge and bottom-edge of plate	10
7.	Minimum clearance between character and left-edge and right-edge of plate	10

Source: Alias (1999)

Peninsular of Malaysia

Sabah

Sarawak

Figure 2.4: Malaysia's car plate specifications endorsed by The Road and Transport Department of Malaysia

Source: The Road and Transport Department Official Portal

2.2 THE EVOLUTION OF RESEARCH IN CAR PLATE RECOGNITION

This section discusses on the evolution of research on car plate recognition systems. The researcher believes that the evolution of the research has started earlier than 1989; however, the discussion was based on the findings from available resources. Table 2.2 shows the summary of some of the findings.

Table 2.2: Summary of researches on car plate recognition

Year	Researcher	Segmentation Technique	Recognition Technique	Recognition Rate
1989	Williams et al	N/A	Syntax forcer	93%
1991	Miyamoto et al	N/A	Mahalanobis distance	More than 99%.
1995	Tindall	Segmentation algorithm	Neural networks	Over 90%.
	Nijhuis et al	Fuzzy c-means clustering algorithm	Fuzzy logics and neural networks	98.51%.
1997	Dragichi	Differential gradient edge detection approach	Neural network	98%
1998	Parisi et al	Finding white areas between columns with higher density of black pixels	Feed forward neural network trained with Block Recursive LS algorithm	90%.
	Lim et al	N/A	Simple character recognition algorithm	More than 90%
1999	Sirithinaphong and Chamnongthai	Horizontal and vertical projection	Four layers BP neural network with supervised learning	92%.
	Yap et al	Blob analysis	Fuzzy ARTMAP neural network	N/A
2000	Naito et al	Geometrical Property (the range of the size of each character)	Template matching	More than 97%.
2001	Emiris and Koulouriotis	N/A	Full Text Recognition option	More than 80%.
2002	Vázquez et al	Labelling process	Multilayer neural network	91.2%.