NUMBER PLATE RECOGNITION SYSTEM USING ARTIFICIAL INTELLIGENCE

MOHD ZA'IM BIN ABDUL A'ALIM

KOLEJ UNIVERSITI TEKNIKAL KEBANGSAAN MALAYSIA

"I admit that to have read this report and it has follow the scope and qualify in Partial Fulfillment of requirements for the Degree of Bachelor of Electronic Engineering (Computer Engineering)"

Signature : 

Supervisor's Name : Encik Abdul Majid Bin Darsono

Date : 5/5/2006

NUMBER PLATE RECOGNITION SYSTEM USING ARTIFICIAL INTELLIGENCE

MOHD ZA'IM BIN ABDUL A'ALIM

This Report is Submitted In Partial Fulfillment Of Requirements For The Bachelor Degree Of Electronic Engineering (Computer Engineering)

Fakulti Kejuruteraan Elektronik Dan Kejuruteraan Komputer
Kolej Universiti Teknikal Kebangsaan Malaysia (KUTKM)

APRIL 2006

DECLARATION

"I admitted that this reports is my own works except for the sentences or phrases that I have states its sources"

Signature	:	
Writer's Name	:	Mohd Za'im Bin Abdul A'Alim
Date	:	05/05/06

ACKNOWLEDGEMENT

Firstly, I would like to say a prayer to Allah the Almighty for always being there for me.

I would like to extend my almost appreciated and gratitude to Mr. Abdul Majid Bin Darsono for his patient guidance and encouragement throughout the implementation of this project.

Special thanks to beloved my family and relatives for their encouragement, understanding and support.

Not to forget, my dear friends especially Safwan, Ram, Kama, Che, Syaman and many other for their assistance and support.

Lastly, my heartfelt appreciation goes to all who have directly or indirectly played parts in the completion of this project

ABSTRACT

This report describes analysis, design and implementation of a system for recognition of number plate. The input to the system is a series of color images of a car, and output consists of the registration number of the number plate. Extraction of the desired information is done in three steps. First, the number plate is extracted from the original image, then the seven characters are isolated, and finally each character is identified. The algorithms were developed using a set of training images to settle all the steps. The final program is capable of extracting the desired information in a high percentage of the test images. This project develops by using MATLAB version 6.5.

ABSTRAK

Laporan ini menerangkan tentang analisis, rekabentuk dan pelaksanaan sistem pengesan nombor pendaftaran kenderaan. Input bagi sistem ini ialah beberapa siri imej kereta dan outputnya ialah nombor pendaftarannya. Keseluruhan pelaksanaan projek ini dapat disiapkan dengan tiga langkah. Pertama, nombor kenderaan akan diekstrakkan dari imej asal, kemudian tujuh aksara nombor kenderaan akan dipisahkan antara satu sama lain dan langkah terakhir ialah mengenalpasti setiap aksara. Algoritma yang telah diuji dengan beberapa siri set imej akan digunakan untuk menyelesaikan kesemua langkah tersebut. Akhir sekali, sistem ini mampu mencapai keputusan yang dikehendaki dengan peratusan yang tinggi. Projek ini menggunakan perisian Matlab versi 6.5.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

It is believed that there are currently more than half a billion cars on the roads worldwide. All those vehicles have their vehicle identification number as their primary identifier.

The vehicle identification number is actually a plate number, which states a legal license to participate in the public traffic. All vehicle world-wide should have its plate number - written on a license plate - mounted onto its body (at least at the back side) and no vehicle without properly mounted, well visible and well readable license plate should run on the roads. To process, sort or analyze data everyone thinks about using computers. If the data is already in the computer most of these tasks are rather easy to be carried out. It is needless to say that the plate number is the most important identification data a computer system should treat when dealing with vehicles.

Suppose a company's security manager would like to have a system that precisely tells at every moment where the cars of the company are: in the garage or out on roads. By registering every single drive-out from and drive-in to the garage, the system could always tell which car is out and which is in.

1.1 Objective

This project is to develop number plate recognition system, which is able to analyze the image of a car given by a camera, locate the registration plate and recognize the number plate of the car

1.2 Scopes of work

The scope of this project is to develop software to analyze number plates. This project is focusing on the processes of recognition of the car number plate, which begin with detecting, and extracting of image region from source images flow till recognition of character presented on the plate number. Image processing techniques such as edge detection, thresholding and resampling have been used to locate and isolate the number plate and the characters. The project was developed by using MATLAB 6.5 software.

1.3 Problem statement

Nowadays, ticket or membership card is commonly used for the entrance at a car park. This is manual system is not flexible because of the time wasted. Number plate recognition system is chosen to overcome this problem. This system, which can recognize number plate that have been register in database and if there are any match with the database, the car allowed to enter. In addition, this system can be extremely useful for the security, airport parking, traffic control and etc.

CHAPTER II

LITERATURE REVIEW

2.1 Number Plate Recognition System

Number Plate Recognition is special type of Optical Character Recognition (OCR) and therefore the definition of number plate recognition should clearly reflect that it is indeed an OCR.

Optical Character Recognition - simple saying - means a software that is able to 'read' the text of scanned documents: *"Optical Character Recognition (OCR) is the method used by a computer to convert scanned in pages of text into electronic text documents."* [1] Specializing the definitions of OCR - and following the accustomed terminology - Number Plate Recognition may be defined as:

"Number Plate Recognition is the method used by a computer to convert digital images of vehicle license plates into electronic text." [1]

"Number Plate Recognition refers to the branch of computer science that involves reading text from digital pictures of vehicle license plates and translating the images into a form that the computer can manipulate." [2]

"Number Plate Recognition converts digitized image of vehicle license plates into editable license number text." [2]

All of these definitions - translated from the definitions of OCR - suggest a computer program, software, a set of algorithms much rather than some hardware devices or integrated systems to be understood under the term 'License Plate Recognition'.

2.2 The Main Steps of the Number Plate Recognition (NPR)

As a first step, the NPR software should localized (located) the possible number plate and then extracted from the image for further processing.

Figure 2.1: Locating the number plate on a digital image of a vehicle

The next step is the character segmentation: after the extracted number plate image is normalised, the individual characters has to be distinguished (segmented) from each other.

Segmentation becomes difficult when the plate is not clear, the characters are touching each other, there are screws or strong light-effects (like shadows) on the plate, etc.

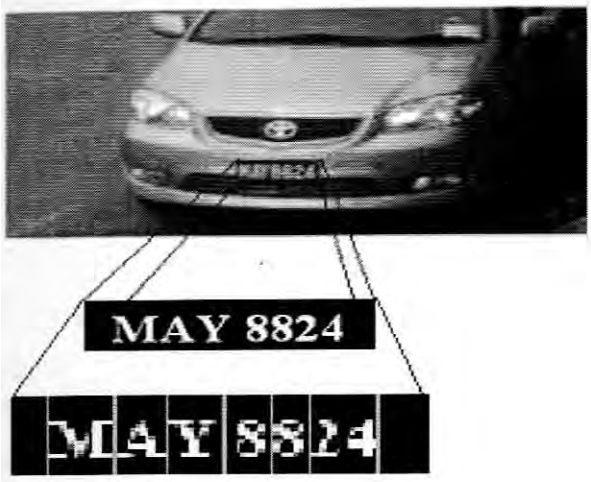

Figure 2.2: Segmenting the characters of the Number plate

When the characters are properly segmented (separated from each other and precisely localized) there is time to invoke the character recognition algorithm for each individual segmented character image. By recognizing all characters after each other, the entire plate text is read.

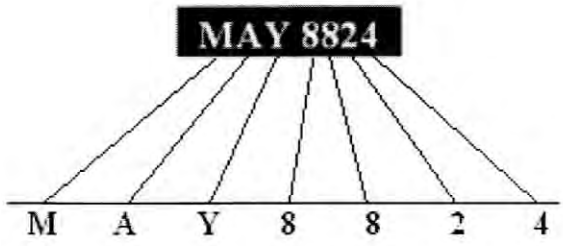

Figure 2.3: Recognising each character

2.3 Type's of the Algorithm

Number plate recognition can be broken down into three main parts:

1. Identifying possible number plates
2. Extracting the characters from the number plate
3. Identifying the characters from the number plate

Each part has their algorithms. This section describes the type's algorithm which can use for this project. Each algorithm has advantage and disadvantage.

2.3.1 Pre-Processing

Pre-processing is carried out on the image to improve the quality of the image so that the main processing on the image becomes easier. This section shows the different pre-processing algorithms that have been encountered while researching number plate recognition.

2.3.1.1 Binarization

Binarization transforms a colour or grey scale image into a black and white image. It is carried out to make the image much easier to work with. For each pixel in the image, its intensity level is examined if it is above a certain threshold level it is turned to white. If it is below the set threshold it is set to black. The threshold level will be some value between 0 and 255.

2.3.1.2 Histogram Equalisation

Histogram equalisation is a function that is used to improve the contrast of an image, which can be used to improve the results of the edge detector. It achieves this by spreading the distribution of grey levels wider and more evenly, ideally to an equal number of pixels per grey level. An ideal number of pixels at each grey level would be calculated using the following formula

$$\frac{\text{rowmax x colmax}}{\text{g\_levels}}$$

where *rowmax x colmax* is the number of pixels in an image and *g_levels* is the number of grey levels over which the spread is required Pixels in the image that have the lowest grey level are allocated a new grey level 0 in the output image. If the resulting grey level 0 has got less than its share of pixels, the next lowest grey level in the original image is allocated to the new output image.

When grey level 0 has gotten its fair share of pixels in the new output image the same process is carried out on the next grey level (i.e. grey level 1), starting with the unallocated pixels that have the lowest grey level in the original image. If the grey level 0 used twice its share of pixles in the new output image then it will have also used up the allocation to grey level 1, in such a case grey level 1 is ignored and allocation moves up to grey level 2. This process is continued until all the grey levels have been allocated to the new output image.

2.3.1.3 Thinning

Thinning is a morphological operation that is used to reduce all black components in a binary image to single pixel wide branches. Thinning does this while preserving the following properties:

- It does not remove end points
- It does not break connectedness
- It does not cause excessive erosion of the region

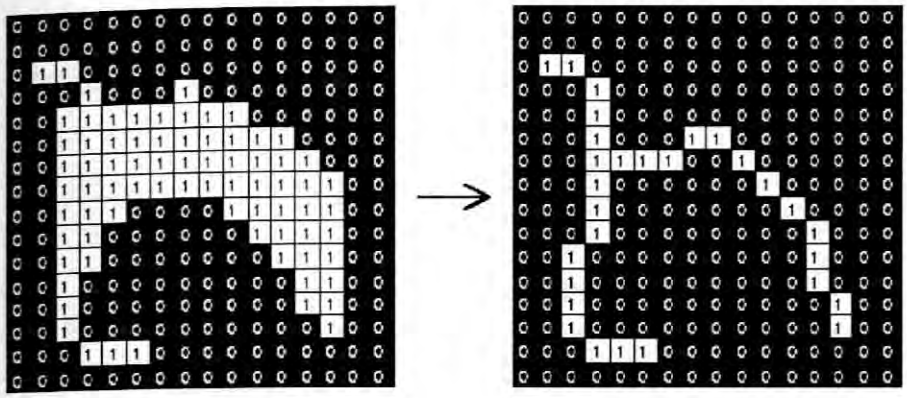

Figure 2.4: The results of a thinning operation on a simple binary image

2.3.2 Identifying Number Plate Candidates

The first objective of number plate recognition is to locate the number plate in the image. The purpose of this section is to identify possible candidate regions of the image in which the number plate might be contained.

2.3.2.1 Region Growing

The basic idea behind region growing is identifying the characteristics of the number plate such as colour and then checking each pixel in the image for the identified characteristics. If a pixel is identified as containing the characteristics of a number plate pixel all its neighbouring pixels are checked to see if they contain the characteristics of a number plate pixel.

If a neighbouring pixel contains the correct characteristics then we say that both pixels belong to the same region. This whole idea is recursively carried out until every pixel in the image has been examined.

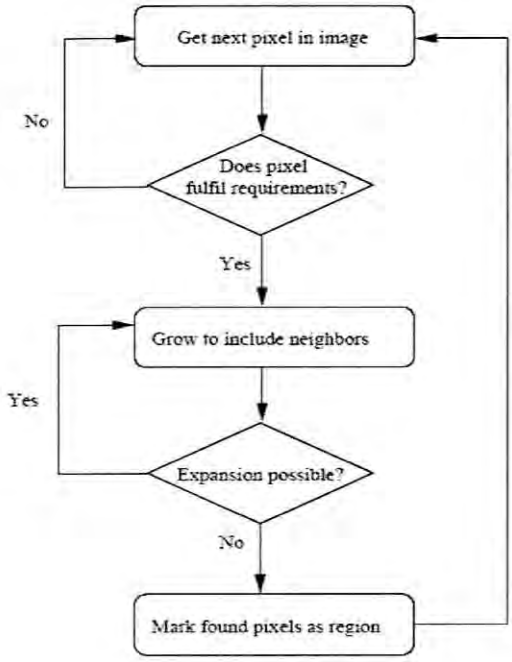

Figure 2.7: The steps of region growing

Thresholding is carried out on the image firstly so that the image to be processed is a binary image (i.e. Binarization). This makes identifying the number plate a lot easier, as the number plate is usually one of the brightest objects in the image as it is white.

The pixels are then recursively examined for regions that meet the criteria. After pixel is examined it is marked, to avoid a region been extracted several times. Region growing does not guarantee that identified regions meet the criteria of being the same dimension or shape as a number plate. This problem can be overcome in two different ways.

Either a region could be transferred into the largest contained rectangle in its area or a region could be transferred into the largest rectangle based on the maximum and minimum co ordinates of the pixels the region contains for both the x axis and y axis. The first method is susceptible to noise so the second method is usually the preferred option.

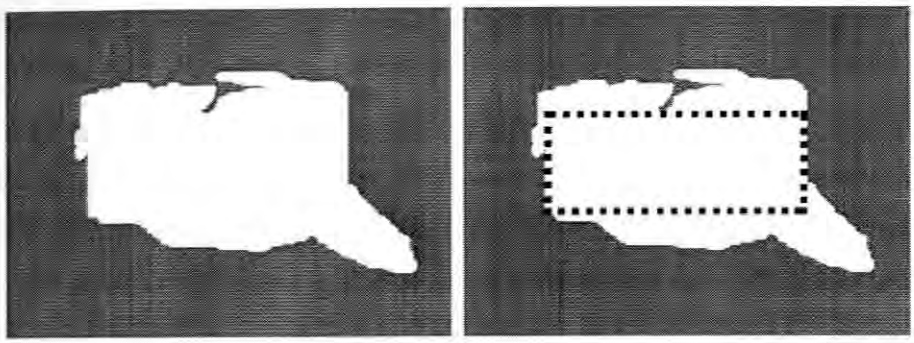

Figure 2.8: Largest Contained Rectangle

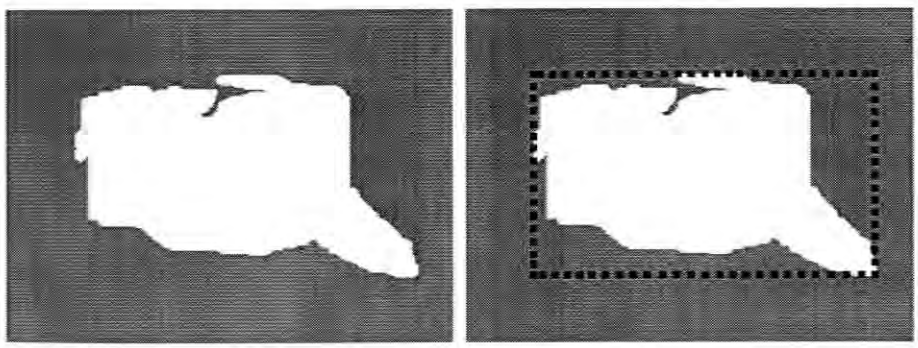

Figure 2.9: Rectangle from maximum and minimum values

Region growing is a fast algorithm (i.e. O(n) algorithm) since each pixel is only examined once. Other advantages are that it extracts candidates with the correct shape, as it does not depend on the size of the region. Using the second method described in the paragraph above to make the identified regions be the same dimension or shape as a number plate makes region growing resistant to noise.

A disadvantage of region growing is that since it is a recursive algorithm it requires a lot of memory usage. Setting the correct threshold for the pre-processing can also cause problems too when trying to create the binary image.

2.3.3 Character Extraction

To simplify the character recognition stage, the characters must first be extracted from the license plate into individual images. There are several approaches that can be taken to extract the characters, some of which are described in the following section.

2.3.3.1 Region Splitting

Region splitting is used to divide part of an image into regions of similar type. This can be applied to the resulting number plate that was extracted from the image.

Region splitting identifies peaks in a grey level histogram and then looks between the peaks for possible threshold values. Regions are identified as containing grey levels between the threshold. This technique relies heavily on the overall histogram giving good guidance as to sensible regions.

No tests were provided with this method so it cannot be concluded how successful this method is.