



Addis Ababa University
School of Graduate Studies
Addis Ababa Institute of Technology

**Design and Implementation of Car Plate Recognition
System for Ethiopian Car Plates**

By: Huda Zuber Ahmed

Addis Ababa
Ethiopia
November, 2011



Addis Ababa University
School of Graduate Studies
Addis Ababa Institute of Technology

**Design and Implementation of Car Plate Recognition
System for Ethiopian Car Plates**

A thesis submitted to the School of Graduate Studies of Addis Ababa University in partial fulfillment of the requirements for the Degree of Master of Science in Computer Engineering

By: Huda Zuber Ahmed
Advisor: Dr. Assefa Dagne

Addis Ababa
Ethiopia
November, 2011



Addis Ababa University
School of Graduate Studies
Addis Ababa Institute of Technology

**Design and Implementation of Car Plate Recognition
System for Ethiopian Car Plates**

By: Huda Zuber Ahmed

Approval by Board of Examiners

<u>Dr. Ing. Getahun Mekuria</u> Chairman, Department of Electrical and Computer engineering	_____ Signature
<u>Dr. Assefa Dagne</u> Advisor	_____ Signature
<u>Dr. Kumudha Raimond</u> Internal Examiner	_____ Signature
<u>Dr. Getachew Alemu</u> External Examiner	_____ Signature

ACKNOWLEDGEMENT

First, I would like to forward my deepest appreciation and thanks to my advisor Dr. Assefa Dagne for his valuable advice and support.

Then, I would like to thank my father Zuber Ahmed, my aunt Nuwera Shafi, Abdurahim Yusuf and the rest of my family for their support, encouragement and ideas. With out you, it would have been difficult for me.

At last, but not least, I would like to thank my class mates Dinkisa Aga, Biniyam Melese and Abadi Tsegay for their encouragement, comments and supports for accomplishing my thesis.

TABLE OF CONTENT

TABLE OF CONTENT.....	5
LIST OF FIGURES	7
LIST OF TABLES.....	8
ABBREVIATIONS	9
ABSTRACT.....	10
CHAPTER 1	11
INTRODUCTION	11
1.1. BACKGROUND	11
1.2. STATEMENT OF THE PROBLEM.....	12
1.4. METHODOLOGY	14
1.5. THESIS SCOPE	14
1.6. CONTRIBUTION AND LIMITATIONS.....	15
1.7. ORGANIZATION OF THE THESIS.....	15
CHAPTER 2	17
REVIEWS ON CAR PLATE RECOGNITION SYSTEM.....	17
2.1. INTRODUCTION.....	17
2.2. LITERATURE REVIEWS	18
2.3. METHODS OF CPR SYSTEM	23
2.4. CONCLUSION.....	29
CHAPTER 3	30
SYSTEM DESIGN AND DEVELOPMENT.....	30
3.1. INTRODUCTION.....	30
3.2. SYSTEM COMPONENTS	30
CHAPTER 4	43
RESULT AND DISCUSSIONS.....	43

4.1.	PLATE REGION DETECTION AND EXTRACTION.....	43
4.2.	OPTICAL CHARACTER RECOGNITION SYSTEM.....	45
4.3.	EVALUATION.....	48
4.4.	SUMMARY	49
CHAPTER 5		50
CONCLUSION AND RECOMMENDATIONS		50
5.1.	CONCLUSION	50
5.2.	RECOMMENDATIONS	51
APPENDICES		52
APPENDIX A: GUI for Car Plate Recognition System.....		52
APPENDIX B: GUI for the Amharic Unicode Converter.....		53
APPENDIX C: Sample Codes.....		54
REFERENCES		58
DECLARATION		62

LIST OF FIGURES

Figure 2.1: Pixel Connectivity (a) 4-connectivity (b) 8-connectivity.....	26
Figure 3.1: Main components of CPR system	30
Figure 3.2: Sample car plate images (a) for Addis Ababa and (b) for Amhara region.....	33
Figure 3.3: Preprocessed image (a) Original image and (b) Image after gray scale conversion and filtering.....	35
Figure 3.4: Components of plate region detection and extraction.....	35
Figure 3.5: Detected and extracted plate region (a) CCA analysis and detected candidate regions (b) The extracted plate region.....	38
Figure 3.6: OCR system components.	39
Figure 3.7: Sample template images.....	41
Figure 4.1: Correctly detected and extracted plate images (a) original image, (b) the binary image and (c) Plate region	44
Figure 4.2: Incorrect detection and extraction (a) original image, (b) The binary image (c) Incorrect region.....	45
Figure A: GUI for the Car Plate Recognition system.....	52
Figure B: GUI for Amharic Text Display.....	53

LIST OF TABLES

Table 3.1: Plate Tasks	31
Table 3.2: Plate specifications	32
Table 4.1: Test results for plate region detection and extraction.....	43
Table 4.2: Test result of the OCR system	46
Table 4.3: CPR Test Result.....	47

ABBREVIATIONS

CPR – Car Plate Recognition

OCR – Optical Character Recognition

CCA – Connected Component Analysis

CCL – Connected Component Labeling

PCA – Principal Component Analysis

LDA – Linear discriminant Analysis

SVM – Support Vector Machine

ANPR – Automatic Number Plate Recognition

ITS – Intelligent Transport System

LPR – License Plate Recognition

HSL – Hue, Saturation,

HSV – Hue, Saturation, Level.

ABSTRACT

During the past few years, Intelligent Transportation Systems (ITSs) have had a wide impact in people's life as their scope is to improve transportation safety and mobility and to enhance productivity through the use of advanced technology. A growing demand for traffic data concerning traffic flow and automatic car identification, led researchers around the world to adopt advanced electronic and computer vision technologies to monitor and control traffic flow. Increasing levels of road traffic needs real time analysis of a moving car in order to extract important data, in this case the car plate number. Also, the incorporation of Optical Character Recognition (OCR) with different applications created the ambition to replicate human tasks with machines.

In this thesis, a car plate recognition system for Ethiopian vehicles is proposed, on the basis of connected component analysis (CCA) for plate region extraction in conjunction with template matching technique for character recognition. First the plate region is extracted and this image is fed to the OCR. The system is simulated in Matlab, and performance is tested on real images. The developed system was tested using 23 test images. The CPR system developed showed an accuracy of 52.63%. Also a separate test was done on the plate region localization and the OCR system. The result of the separate tests showed accuracy of 82.6% for the plate region localization and 82.11% for the OCR. It is observed that the results are mainly influenced by the quality of the input image and distance between the camera and the car. The experimental results showed that by adjusting the influencing factors, the system can perform in a good way.

Key words: Optical Character Recognition (OCR), Connected Component Analysis (CCA), Plate region extraction.

CHAPTER 1

INTRODUCTION

1.1. BACKGROUND

Massive integration of information technologies into all aspects of modern life caused demand for processing vehicles as conceptual resources in information systems [15]. With the rapid development of public transportation system, automatic identification of vehicles has played an important role in numerous applications such as managing park facilities, controlling traffic volume, detecting stolen vehicles, automatic toll connection, and so on.

The Automatic Number Plate Recognition (ANPR) was invented in 1976 at the Police Scientific Development Branch in the UK. However, it gained much interest during the last decade along with the improvement of digital camera and the increase in computational capacity [1].

An automatic car (license) plate recognition system plays an important role in a modern traffic system or intelligent transport system (ITS). Car plate recognition system is a very important area which encompasses the application of OCR for detecting and identification of vehicles for different purposes. It also includes image processing concepts such as feature extraction and representation techniques. Car plate recognition system can be used in several areas such as police law enforcements, parking area control, security management, driving speed law enforcement and so on.

OCR is a mechanism for translating images of hand written, type written or printed text into machine encoded format. It is a field of research in pattern recognition, computer vision and artificial intelligence areas. It allows a machine to automatically recognize characters using optical recognition mechanism.

OCR can also be used for converting a text image to editable text and text mining. OCR makes it possible to edit the text, search for a word or phrase, store it more compactly, display or print a

copy free of scanning artifacts, and apply techniques such as machine translation, text-to-speech and text mining to it. OCR is a field of research in pattern recognition, artificial intelligence and computer vision.

In Ethiopia, the traffic control is currently being managed by traffic polices. This process is often very much influenced by the physical limitations faced by the traffic polices. For example, a traffic police cannot register the plate if the vehicle is moving fast or if the weather conditions is bad (rain, fog). In such situations, the car plate recognition system can perform the task in a better fashion. That is, it can help a human operator and improve the overall quality of a service.

Number plate recognition system comprises two common stages;

- 1. Detecting and extracting the plate region:** It is the process of localizing the plate region from the whole image. It involves several image processing concepts such as noise filtering, component analysis, color conversion, segmentation and others. Plate localization strategies that involve smaller computational resources for the purpose of producing fastest algorithm are an active research area to achieve real-time performance [16]. This is because the recognition accuracy is directly influenced by the localization result.
- 2. Recognition of characters in the image:** After the plate region is extracted, the individual characters on the plate should be extracted and recognized. The main processing in this stage is finding and recognizing characters (numbers and alphabets) on the plate.

1.2. STATEMENT OF THE PROBLEM

Car plate detection and recognition is a form of OCR technique used to identify a car by its license plate with minimal human involvement. This technique has got several applications in traffic monitoring and evaluation schemes.

Some of the obvious implementations of this image processing technology are identification of traffic rule violators, automobile tracking, security monitoring, etc. Especially in identifying traffic rule violators; the technique contributes immensely due to the fact that usually identification is done manually. There are obvious draw backs associated with undertaking tasks manually such as lack of efficiency, problem of record keeping, unease of retrieval, etc. The electronic processing of the image gathered by Closed Circuit TV (CCTV) cameras would result in a digitalized form of the license plate ready for further processing be it storage, matching, history gathering and so on. OCR is also used for a wide range of character recognition applications which includes from transformation of anything humanly readable to machine manipulatable representation. This eases the barrier of interface between human and machine to great extent by saving time and cost in converting character on handwritten or images into electronic format.

In general the main motivation of this research is to create a system that requires less human involvement for the traffic control process to increase the manageability of the traffic system.

1.3. OBJECTIVE

1.3.1. General objective

Given an image of a car with a plate, the general objective is to develop an offline system that is capable of localizing the car plate from the image and an OCR system, which is capable of recognizing the characters and numbers on the plate.

1.3.2. Specific objective

The specific objectives are

- Reviewing different techniques and algorithms for developing the CPR system.
- Selecting the appropriate mechanism for localization and detection of plate region from the image.

- Selecting an appropriate recognition algorithm to develop an OCR system.
- Developing plate region detection and localization system for localizing the plate region from the image.
- Developing OCR system which is capable of recognizing the special Amharic characters on the plates (Amharic characters that are used for specifying regions).
- To test the performance of the developed system using the test images.

1.4. METHODOLOGY

Literature review and assessment: A literature analysis and theoretical assessment of existing systems is done. Different existing systems are studied and compared in order to find out a way to solve the problem.

System Design: After choosing the relevant techniques, the system was designed using the selected technique for each stage.

Performance evaluation: In order to check the accuracy of the system, it is tested using the data collected. Some of the images are obtained by capturing real time images and some are obtained from internet. Based on the test result, some modifications are done on the system.

1.5. THESIS SCOPE

The scope of this thesis lies in identifying Ethiopian standard car plates. That is, identifying the number on the plates and the region. Secondly, it is designed to work for identifying a single plate at a time. That is, it is an offline system that is capable of detecting a single car plate in an image.

1.6. CONTRIBUTION AND LIMITATIONS

The ambition to replace human tasks by machines led to invention of different intelligent systems that are capable of processing, analyzing and manipulating information alone. Some of these systems are speech recognition, image processing, computer vision and other. Car plate recognition system is one of the image processing applications. It can be implemented for several purposes such as traffic control, security management, police law enforcement and others.

So far, there are many car plate recognition systems which are implemented for several purposes in many countries. But in Ethiopia, there is no system so far designed for the Ethiopian car plate specification. Additionally, some researches are done on Amharic OCR. From this view, the main contribution of this thesis is to develop a system that works for Ethiopian car plate standard and a system which incorporates the use of Amharic OCR for recognizing the special Amharic characters on the plate. This can be implemented for several areas such as Ethiopian traffic department, police law enforcement department, security management, parking management and others.

The main limitation of this study was acquiring the sample plate images. Few samples were found on internet but most of the samples were taken manually. In doing this, it was difficult to take photo of plate images because it was some how difficult to convince the people the need to take the photo. Also, most of the test images are Addis Ababa region plates because there was difficulty to obtain the regional plates.

1.7. ORGANIZATION OF THE THESIS

The remaining chapters of the thesis are organized as follows. The second chapter discusses the CPR system and the different works done on CPR. Different techniques of CPR and OCR systems are also discussed as reviewed in different literatures. The basic principles and methods of OCR system are also included.

The third chapter presents the techniques used to develop the system. The fourth chapter is about the experimentation result of the developed CPR system.

Based on the experimentation results, the conclusion and recommendations of the study are discussed in the final chapter.

CHAPTER 2

REVIEWS ON CAR PLATE RECOGNITION SYSTEM

2.1. INTRODUCTION

Human beings are predominantly visual creatures: we rely heavily on our vision to make sense of the world around us. Humans have evolved very precise visual skills: we can identify a face in an instant, differentiate colors, and process a large amount of visual information very quickly and others. The ambition to create machine that are capable of recognizing information from images led to the development in the field of image processing. Image processing involves changing the nature of an image in order to either improve its pictorial information for human interpretation, or render it more suitable for autonomous machine perception. It has wide applications in computer vision, artificial intelligence, detection systems, monitoring systems and others.

Solutions to the problems of image processing require extensive experimentations and a frequent updating of the techniques by studying and comparing accuracy and performance efficiency of different algorithms. One of the applications of image processing is recognition systems. In recent years, recognition and detection systems are having a great impact in our day to day activities. It is becoming as the basic need in different applications such as banking tax payment, security management, monitoring and control applications and other applications. Car plate recognition system is one of the many applications.

CPR system is a mass surveillance method that uses optical recognition systems on images to read the plate on vehicles. It is an image processing technology, which uses number (license) plate to identify the vehicle. A CPR system is able to recognize vehicles automatically and therefore useful for many applications such as portal controlling, traffic monitoring, stolen car detection, and etc [17]. During the past few years, intelligent transportation systems (ITSs) have had a wide impact in people's life as their scope is to improve transportation safety and mobility and to enhance productivity through the use of advanced technologies [8]. Due to this, CPR has

been an interesting area for researchers and computer vision experts. It is also known by names such as: Automatic number plate recognition (ANPR), License plate recognition (LPR) and others.

The main components of a CPR system are; the plate region localization and detection system and character segmentation and recognition system (optical character recognition). The plate localization and detection system is responsible for finding the license plate from the image. It uses different features and properties of the plate in order to select the plate region from the other candidate regions. Some of the properties that it makes use of are: morphological properties, shape, color, brightness and orientation (Horizontal and vertical). Number plate localization strategies that involve smaller computational resources for the purpose of producing fastest algorithm are an active research area especially to achieve real-time performance [16]. This is due to the fact that, the recognition accuracy and efficiency is directly influenced by the localization strategy. The second component, the recognition system, is used to segment the character for recognition. It aims in finding the individual character on the plate and extraction for detection purpose.

There are many works done on CPR systems. In this chapter, different research and works done on CPR systems will be discussed and reviewed.

2.2. LITERATURE REVIEWS

The car plate recognition system has been a popular field of research for several years. It is a system capable of recognizing car plates automatically. There are different works done on CPR using different approaches. There are many techniques that have been introduced by different researchers to localize the number plate from an image and recognize the characters.

2.2.1. PLATE LOCALIZATION AND DETECTION

The plate localization and detection is first phase for CPR system. It aims at extracting the plate from the image. It directly influences the accuracy of the CPR system. It includes several image

processing concepts. The general system developed for plate localization and detection comprise of the following main processes: Preprocessing and Plate location detection and extraction. The main objectives and the works done on each phase will discussed on the below subtopics.

1. Preprocessing

Original images often suffer from various types of degradation such as noises and low contrast. Therefore, original images need to be preprocessed [2]. Image preprocessing step is taken before major processing tasks. The aim of this stage is to perform some basic tasks in order to render the resulting image more suitable for the job to follow. Preprocessing stage incorporated several image processing concepts such as noise removal, filtering, contrast enhancement techniques, color conversion and others.

In most researches, the first step of preprocessing stage is conversion of the input image, which is a colored image, into gray scale image. A gray scale image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information.

In [2], W. Wen et al., first converted the original image to gray scale image. Due to the impact of different light, the license plate is not prominent against the background. Due to this, gray balancing is used to improve the image contrast. Finally, to filter out the small particle noises, Gaussian Filter is chosen.

J.R. Parker et al. [5] converted the input image (color image) into gray scale image. Then a median filter (5x5) is applied on the gray scale image in order to remove the noise, while preserving the sharpness of the image.

2. Plate Location Detection and Extraction

Due to the variations of brightness, size, angle and quality of the plate image, the location algorithms always degrade dramatically [2]. The fact that the recognition system is influenced by the localization method increased the motivation of researchers to work on different algorithms.

A number of methods have been used for plate region detection. These include morphological operations, edge statistics (vertical and horizontal edge detection), color based operations, transforms (Wavelet transform, Hough transform) and others.

According to [1], a color search technique is implemented to extract the likelihood region of interest in an image. In a similar way, in [10], Jia LI and Mei XIE used the HSI (Hue, Saturation, Intensity) color model values and the histogram analysis result to extract the plate region. Also, in [3], Xianmin Wei, developed a color based technique. First, the input image is roughly segmented based on the color information, and some regions with the color of car-plate are obtained. Then the false regions in the result image are eliminated. At last, projection method is used to get the accurate position of the plate.

Also edge detection technique, especially, vertical edge detection technique is also implemented to detect the regions. An edge, in general, is defined as a local discontinuity in the pixel values which exceeds a given threshold. Since the pixels level of the plate region is different from the background, this attribute can be used for edge detection. According to [6, 7], vertical edge detection algorithm (VEDA) is used to extract the candidate region on the image. After edge detection, we need to identify the region of interest from the candidate regions. In [7], Lin Luo et al. used connected component analysis (CCA) to locate the plate region, while in [6], statistical operations (that is, analyzing the lines in each row and column using the pixel property) is implemented.

In other approaches, morphological based techniques (using the forms and structure of the plate) are used [11, 12]. In these approaches, characteristic features such as license plate width and height, character height and spacing are considered for defining structural elements for morphological operations. Similarly, a rule based candidate filtering is proposed in [16]. In this method, the algorithm comprises of several rules that depend on dimension and compactness analysis, ratio and angle analysis, horizontal vertical analysis and candidate verification.

Sliding concentric window (SCW) technique is also used [8, 9]. In SCW, the standard deviation values of neighboring pixels are analyzed for the extracting purpose.

In general, there is a large volume of research work done on plate extraction process. Each technique contributed in the advancement of the field. This can imply to the fact that it is an active area of research because it highly contributes in the advancement of the CPR technology.

2.2.2. RECOGNITION SYSTEM (CHARACTER RECOGNITION SYSTEM)

Character recognition system is the system, which is used to extract the characters on the plate. It segments the individual character on the image and each character is recognized individually. It depends on the accuracy of the plate localization process and similarly the outcome of the CPR system depends on the efficiency of the recognition system. The main tasks in each stage are: Segmenting the individual characters and recognition process. In most researches, the segmentation process is achieved by horizontal and vertical pixels processing while for the recognition system several techniques such as correlation method, neural networks, Hough transforms and others are used.

Muhammad T. Qadri and Muhammad Asif [1] developed an OCR to recognize the characters. Their OCR used row and column scanning to segment the individual characters and a correlation technique to match each character. In another paper [8], the researchers used SCW technique to segment the characters and Neural Network for the recognition process. Also in [11], the recognition of the contents of the License Plate is performed using cross correlation followed by Neural Network.

2.2.3. OPTICAL CHARACTER RECOGNITION SYSTEMS

Machine replication of human functions, like reading, is an ancient dream. However, over the last five decades, machine reading has grown from a dream to reality. OCR has become one of the most successful applications of technology in the field of pattern recognition and artificial

intelligence. Many commercial systems for performing OCR exist for a variety of applications, although the machines are still not able to compete with human reading capabilities [18].

OCR is used for a wide range of character recognition applications which includes anything from transformation of anything humanly readable to machine manipulatable representation. OCR aims at enabling computers to recognize optical symbols without human intervention [22]. It has been an active research in the field of image processing and computer vision. OCR systems are typically configured for a specific script and/or a domain of interest.

Many research works have already been done through last two and half decades regarding different aspects of OCR system. The research initially started with printed text segmentation and recognition and later the works have been extended to segmentation and recognition of handwritten texts also. The main application of OCR was initially to recognize scanned document images. But with the progress of automation, different technical areas have been evolved, which require expertised OCR system. The past research works were mainly focused on scanned document images [14].

According to [1], in order to segment the individual characters, horizontal and vertical scanning of the pixels is used. Row scanning for line separation and column scanning for character extraction was used. For the classification, a simple correlation technique to match the individual characters is implemented.

For text segmentation, Hough transform method is also used. In [14], the researchers implemented a Hough transform based technique for line and word segmentation from digitized images after preprocessing the image and edge detection process.

Artificial neural networks are also used in classification process for OCR development. According to [23], the characters were segmented by using vertical and horizontal projections. The moments of the character pixels about the center of gravity are used as features for feature extraction process. These features yielded a set of descriptors that are invariant to scaling, translation, and rotation. And finally, an artificial neural network classifier is used for classifying

the characters. They used a multilayer perceptron (MLP) layer with back propagation learning. Similarly, in [8], a probabilistic neural network (PNN) was used for identifying the characters (alpha-numeric characters).

In [13], support vector machine (SVM) classifier is used. Also a two stage feature extraction method was developed, that is, PCA followed by LDA for optimal discriminant feature extraction.

In [24], statistical algorithm was used to identify the characters. After the characters are normalized in size and orientation, it is compared with a set of templates for matching.

2.3. METHODS OF CPR SYSTEM

The main processes of CPR system consists two main stages: Plate region detection and the character recognition. The processes involved in each stage will be discussed.

2.3.1. SEGMENTATION

Segmentation is the process of subdividing an image into constituent parts, or isolating certain aspects of an image such as finding lines, circles or particular shapes in an image. It refers to the operation of partitioning an image into component parts, or into separate objects. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. In general, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

There are different approaches to perform image segmentation with two main categories. These categories are thresholding and edge detection.

1. Thresholding

Thresholding provides an easy and convenient way to perform image segmentation based on different intensities or colors in the foreground and background regions of an image. It segments the image into foreground and background region by using a threshold value. The main challenge in this approach is selecting the optimum threshold value. There are two types of thresholding techniques. These are: single thresholding and double thresholding. In single thresholding, a single threshold value is selected and the pixels values greater than the threshold are segmented into one region and the values less than the threshold in another region. In double thresholding, we choose two threshold values. In general, both thresholding methods can be summarized as:

For single thresholding:

A pixel becomes $\begin{cases} \text{white if its gray level is } > T \\ \text{black if its gray level is } \leq T \end{cases}$

And for double thresholding:

A pixel becomes $\begin{cases} \text{white if its gray level is between } T1 \text{ and } T2 \\ \text{black if its gray level is otherwise} \end{cases}$

Where T , $T1$ and $T2$ are the threshold values.

The main challenge in using thresholding method is finding a good threshold value. One way to choose threshold is by using visual inspection of the image histogram. Another method is choosing the threshold by trial and error method, picking different thresholds until one is found that produces a good result [19]. Also we can use the average value of the gray levels in the image as a threshold.

2. Edge detection

An edge may be loosely defined as a local discontinuity in the pixel values which exceeds a given threshold or it is an observable difference in pixel values. Edges occur on the boundary between

two different regions in an image. Based on this, edge detection is segmenting the image using the difference in pixel values. The purpose of detecting edges is to identify points at which image brightness changes and capture important events and changes in the image. There are a number of algorithms for this, but these may be classified as derivative based where the algorithm takes first or second derivative on each pixel or gradient based where a gradient of consecutive pixels is taken in x and y direction [26]. Some of edge detection methods are gradient-magnitude methods, Sobel, Prewitt, Roberts, next zero-crossings, Laplacian and Canny method.

3. Connected Component Analysis (CCA)

CCA is the process of analyzing and segmenting connected components in a binary image. The goal of the connected component analysis is to detect connected region or object. Each maximal region of connected pixels is called connected component. The set of connected components are used to partition an image into segments. Once region boundaries have been detected, it is useful to extract regions. Connected regions are extracted using the connected component labeling process. Connected component labeling (CCL) is the process of assigning a unique label to each maximally connected region of pixels [28]. It is used to identify objects in binary image. Extracting and labeling of various disjoint and connected components in an image is central to many image analysis applications. Determining which pixels create a connected component depends on how pixel connectivity is defined. There are pixel connectivity standards for two-dimensional and three-dimensional objects. For two dimensional objects, there are 4-connectivity and 8-connectivity while for three-dimensional there are 6-connectivity, 18-connectivity and 26-connectivity. The type of connectivity chosen affects the number of objects found in an image and the boundaries of those objects.

In 4-connectivity, pixels are connected if their edges touch. This means that pair of adjoining pixels is part of the same object only if they are both on and are connected along the horizontal or vertical direction. In 8-connectivity, pixels are connected if their edges or corners touch. This means that if two adjoining pixels are on, they are part of the same object, regardless of whether

they are connected along the horizontal, vertical, or diagonal direction. Figure 2.1 shows the connectivity technique for 4-connectivity and 8-connectivity.



Figure 2.1: Pixel Connectivity (a) 4-connectivity (b) 8-connectivity

CCL scans an image and groups its pixels into components based on pixel connectivity, *i.e.* all pixels in a connected component share similar pixel intensity values. Once all groups have been determined, each pixel is labeled according to the component it was assigned to. It works by scanning an image, pixel-by-pixel (from top to bottom and left to right) in order to identify connected pixel regions, *i.e.* regions of adjacent pixels which share the same set of intensity values V (for a binary image $V=1$). It can be applied on binary or gray scale images using different connectivity measures. For 8-connectivity, the CCL operator scans the image by moving along a row until it comes to a point p (where p denotes the pixel to be labeled at any stage in the scanning process) for which $V=1$. When this is true, it examines the four neighbors of p which have already been encountered in the scan (*i.e.* the neighbors to the left of p , above it, and the two upper diagonal terms). Based on this information, the labeling of p occurs as follows:

- If all four neighbors are 0, assign a new label to p , else
- If only one neighbor has $V=1$, assign its label to p , else
- If more than one of the neighbors have $V=1$, assign one of the labels to p and make a note of the equivalences.

After completing the scan, the equivalent label pairs are sorted into equivalence classes and a unique label is assigned to each class. As a final step, a second scan is made through the image,

during which each label is replaced by the label assigned to its equivalence classes. Similar procedure is followed for 4-connectivity.

2.3.2. FEATURE EXTRACTION

To recognize or classify an object in an image, one must first extract some features out of the image, and then use these features inside a pattern classifier to obtain the final class [20]. Feature extraction is the process of capturing essential characteristics [18]. It is the process of transforming the input data into set of features. Feature extraction reduces the amount of data by extracting the relevant information. The main challenge in feature extraction is to identify the characteristic features that can form a good representation of the object, so as to discriminate across the object category. The criteria for selecting a good feature are; features should be easily computable, preferably independent of translation and size, and chosen so that they do not replicate each other.

Feature extraction (or detection) aims to locate significant feature regions on images depending on their intrinsic characteristics and applications. These regions can be defined in global or local neighborhood and distinguished by shapes, textures, sizes, intensities, statistical properties, and so on. Local feature extraction methods are divided into intensity based and structure based. Intensity-based methods analyze local intensity patterns to find regions that satisfy desired uniqueness or stability criteria. Structure-based methods detect image structures such as edges, lines, corners, circles, ellipses, and so on [20]. The most popular feature extraction techniques are principal component analysis (PCA) and linear discriminant analysis (LDA).

PCA is known as the best data representation in the least-square sense for classical recognition. It is commonly used to reduce the dimensionality of images and retain most of information. The central idea behind PCA is to find an orthonormal set of axes pointing at the direction of maximum covariance in the data. It is often used in representing facial images. The idea is to find the orthonormal basis vectors, or the eigenvectors, of the covariance matrix of a set of images, with each image treated as a single point in a high-dimensional space. LDA is used to perform

the feature extraction task and to classify samples of unknown classes based on training samples with known classes. It provides a linear transformation of k -dimensional samples (or feature vectors in our experiments) into an m -dimensional space ($m < k$), so that samples belonging to the same class are close together, but samples from different classes are far apart from each other. This method maximizes the ratio of between-class variance to within-class variance in any data set; thereby, the theoretical maximum separation in the linear sense will be guaranteed. Since LDA seeks directions that are efficient for discrimination, it is the optimal classifier for distinguishing classes that are Gaussian distribution and have equal covariance matrices [20].

2.3.3. CLASSIFICATION

Classification refers to the process of assigning a given input data into one of a given number of categories or it is the process of identifying each character and assigning it to the correct character class. An algorithm that implements classification, especially in a concrete implementation, is known as a classifier. Classification compares feature vectors to the various models and find the closest match.

Classification is the main task in recognizing characters. There are several classification techniques. Some of the techniques are template matching, neural networks, statistical approach, structural (syntactical) approach and others.

Traditional statistical classification procedures such as discriminant analysis are built on the Bayesian decision theory. In these procedures, an underlying probability model must be assumed in order to calculate the posterior probability upon which the classification decision is made. One major limitation of the statistical models is that they work well only when the underlying assumptions are satisfied. The effectiveness of these methods depends to a large extent on the various assumptions or conditions under which the models are developed. Users must have a good knowledge of both data properties and model capabilities before the models can be successfully applied [25].

Template matching is a technique in digital image processing for finding small parts of an image which match a template image. There are two approaches for template matching: feature based and template based matching. In feature based matching, features for search, such as corners and edges to find the best matching location of the template and template image for correlating the result of the primary search. In template matching, classification is performed by comparing an input character image with a set of stored templates (or prototypes) from each character class. Each comparison results in a similarity measure between the input character and the template. The similarity measure, often a correlation, may be optimized based on the available training set. One measure increases the amount of similarity when a pixel in the observed character is identical to the same pixel in the template image. If the pixels differ the measure of similarity may be decreased. After all templates have been compared with the observed character image, the character's identity is assigned as the identity of the most similar template [21]

2.4. CONCLUSION

In general, the car plate recognition system has been an active area of research for several decades. It has contributed for the development of intelligent transport system (ITS). It is widely used in several applications such as portal controlling, traffic monitoring, security monitoring, stolen car detection etc. Yet, the system still faces several challenges such as performance degradation due to several factors such as image quality and environmental conditions. The major challenges in any CPR system are the plate localization technique and the recognition (OCR) system. Also OCR systems are becoming as an essential need in recognition and detection purpose. They have several applications in different aspects of our daily activities such as preserving historical documents, converting type written, hand written, or printed text into machine encoded format, in recognition of text documents etc. In both areas, a lot of progress has been achieved but there are still many open challenges to overcome.

CHAPTER 3

SYSTEM DESIGN AND DEVELOPMENT

3.1. INTRODUCTION

This chapter discusses about the basic algorithm used to develop the car plate recognition system. It can be divided into two main parts; plate region extraction system and the OCR system. It also includes the techniques, design prototypes and the parameters for developing this study.

3.2. SYSTEM COMPONENTS

The over all architecture of the car plate recognition system developed in this study is as follows.



Figure 3.1: Main components of CPR system

In each component, there are many processes and algorithms. In the following sections, the details development specification and algorithms used will be discussed.

3.2.1. PREPROCESSING

In this stage, preprocessing tasks are carried out before the major processing in order to render the resulting image more suitable for the job to follow. The main aim of this stage is to improve the quality of the image for the plate detection and extraction process.

The main tasks in this phase are; converting the input RGB image into gray scale image and filtering process.

The input image for this stage is RGB image. The image contains a car plate and the body of the car. The input images for the conversion stage are Ethiopian car plate images. The Ethiopian car plate's standards have their own specifications and tasks. There are 10 regional plates. These regions are Addis Ababa region, Afar region, Amhara region, Benshangul Gumuz, region, Gambella region, Harar region, Oromiya region, Somali region, Southern Nations and Nationality People (SNNP) region and Tigray region. The tasks of the plates are recognized based on code and color specification. The following tables give a summary on the plate specifications and tasks.

The background color for all the plates is white except for the police plate (which is yellow). The plates come in different size as 460X110mm, 520X110mm, 280X200mm and 340X200mm area. Some of the sample plates are shown in Figure 3.2.

Using this image, the preprocessing stage starts processing the image for the plate region detection and extraction system. The first process is gray scale conversion followed by filtering process to remove the noise.

Table 3.1: Plate Tasks

Code Number	Task	Character Color
1	Taxi, public transport	Red
2	Private cars	Blue
3	Commercial vehicles	Green
4	Higher authority/ Administration vehicles	Black
5	Other Organizations	Orange

Table 3.2: Plate specifications

	Region	English Character	Amharic Character
1	Addis Ababa	AA	አአ
2	Afar	AF	አፋ
3	Amhara	AM	አማ
4	Benshangul Gumuz	BG	ቤጉ
5	Gambella	GM	ጋም
6	Harar	HR	ሐረ
7	Oromiya	OR	ኦሮ
8	Somali	SM	ሶማ
9	Southern Nations and Nationality People (SNNP)	SP	ደሕ
10	Tigray	TG	ጉግ
Others			
	Core Diplomat	CD	ኮዲ
	Police		ፖሊስ
	United Nations	UN	ተመ
	Africa Union	AU	አሕ



(a)



(b)

Figure 3.2: Sample car plate images (a) for Addis Ababa and (b) for Amhara region.

I. GRAY SCALE CONVERSION

Gray scale conversion is the process of converting the true color image (RGB) to the gray scale (intensity) image. In this phase, the Matlab function *rgb2gray()* is used. It converts the true color (RGB) image to intensity image using the formula:

$$\text{Gray image} = 0.2989 * R + 0.5870 * G + 0.1140 * B$$

Where R, G and B are the red, green and blue color components.

Converting the RGB image into gray scale (intensity) image is preferred because:

1. It is more efficient for manipulation and processing.
2. When working with RGB or other color coordinate images such as HSL/HSV, it is frequently necessary to access one plane at a time. But with gray scale images, we have only one plane, the intensity value (that ranges from 0-255). So due to this it reduces access time, which indirectly decreases the computational time.

II. FILTERING

After the RGB image is converted to intensity image, the image may suffer some degradation. Due to this, we need to remove the noise. For this study case, Median filtering is chosen. Median filter is a non-linear filters used to reduce noise. It is a very widely used filter because it preserves edges while removing noise.

In this stage, a 3X3 median filter is used to filter the gray image in order to improve the result of the later processing. The final result of this stage is a filtered gray image. This image is fed to the next component for plate region detection and extraction. Sample image after gray scale conversion and filtering is shown in figure 3.3.



(a)



(b)

Figure 3.3: Preprocessed image (a) Original image and (b) Image after gray scale conversion and filtering

3.2.2. CAR PLATE REGION DETECTION AND EXTRACTION

The main aim of this stage is to find and isolate the plate region on the image. The main processing in this stage is shown in Figure 3.4.



Figure 3.4: Components of plate region detection and extraction

I. SEGMENTATION

In image processing, segmentation is the process of partitioning an image into several segments or partitions. Segmentation is used for locating objects and boundaries in an image. It is also useful to change the representation of an image for easy analysis and manipulation. For example, we can convert an intensity image to binary image using segmentation technique.

There are different types of segmentation techniques. From these methods, thresholding segmentation is selected. Thresholding segmentation involves the process of obtaining a binary image based on the brightness of the pixels, that is, pixels brighter than a specified brightness level are taken as 1 and the rest are left 0 according to the value assignment. This method is chosen because it is a simple technique and it provides an easy and convenient way to perform image segmentation based on different intensities or colors in the foreground and background regions of an image. The main challenge in this technique is selecting the threshold value that will give a good segmentation result. The technique for choosing the threshold value for this system was based on average value of the intensity level. As a starting threshold value, the mid-point value of the gray scale was taken. Then through experiment, the threshold value which resulted in a good segmentation result was found. The input is the gray scale image obtained from the previous stage (preprocessing stage). The general segmentation formula is:

For all pixels with x and y coordinates:

$$b(x, y) = \begin{cases} 1 & \text{if } g(x, y) > 140 \\ 0 & \text{if } g(x, y) \leq 140 \end{cases}$$

Where $b(x, y)$ is the segmented binary image and $g(x, y)$ is the input gray scale image.

The final output of this phase is a binary image. The binary image is filtered after segmentation to reduce the noise resulting from the segmentation process.

II. CONNECTED COMPONENT ANALYSIS (CCA)

Connected component analysis is the process of analyzing connected components in a binary image. Connected component labeling is the process of detecting and labeling connected regions in a binary image. Once region boundaries have been detected, it is often useful to extract regions which belong to the same region. Any set of pixels which is not separated by a boundary is called connected. Each region of connected pixels is a connected component and the set of connected components partitions an image into segments.

Before starting the CCA, we need to label the binary image. Regions with the same label can be segmented as a single object. The Matlab function *bwlabel* () is used to label the connected components in the binary image. Connected component analysis is used to analyze the regions independently and measure the property of image region. Based on this, pixels belonging to the same region are segmented as an object.

Since the plate region is a rectangular shape, the final outputs of this phase are rectangular shaped regions. From the candidate regions, the plate region is selected using a selection process.

III. SELECTION OF PLATE REGION

After the candidate regions are extracted, the plate region should be selected from those regions. In order to identify the regions, certain features, which are unique to plate, are used. These features include the rectangular shape, the size of the plate and the ratio of the white to black pixels. Sample result image of this stage is shown in figure 3.5.



(a)



(b)

Figure 3.5: Detected and extracted plate region (a) CCA analysis and detected candidate regions (b) The extracted plate region.

3.2.3. OPTICAL CHARACTER RECOGNITION SYSTEM (OCR)

The extraction process in a way distills the image to draw out the license plate part of the image. Once the license plate is extracted, the next process would be to make sense of the image, which is, converting it to editable text. Since we are dealing with Ethiopian car license plates, the variables we need to recognize vary from numerical characters to English alphabets and most importantly Amharic characters.

After the plate region is detected and extracted, the only task left is to recognize the characters on the plate image. This can be done using OCR. The general OCR components are shown in figure 3.6.

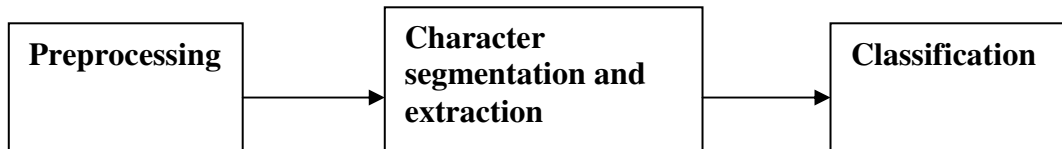


Figure 3.6: OCR system components.

I. PREPROCESSING

In this system, the preprocessing phase is simply improving the image / the extracted car plate image. The input for this phase is a binary plate image, which is obtained from the plate region detection and extraction phase. Noise is reduced and the output of this phase is also a binary image which will be the input for the segmentation and extraction process.

II. CHARACTER SEGMENTATION AND EXTRACTION

The main objective of this phase is to find the individual characters (numbers and alphabets) and extracting the characters on the plate image. For segmentation and extraction process, CCA is also used in this case. Connected region with the same label can be assumed as a single character. Using this approach, the individual characters are segmented and extracted.

First the binary image is labeled using the function $bwlabel()$. The region with the same label is assumed as an object. Based on this concept, a single character is segmented as an object and extracted as a character. The final output of this region is the set of characters on the car plate images, alphabets and numbers. These extracted characters will be passed to the classifier for classification and recognition task. Some regions, which are not character may be segmented. In order to avoid this kind of segmentation, we set a threshold value on the size of the characters. Pixels less than the threshold pixel will not be segmented. This approach reduces segmentation of

unwanted regions and reduces the task on the classifier by minimizing the number of comparison process.

III. CLASSIFICATION

After the individual characters are extracted, the characters should be identified and assigned to the correct character class. This is done by comparing the feature vector to the various models and finding the closest match. For this purpose, template matching technique is used. In template matching technique, the input character is compared with the set of template images from each character class. Based on the similarity value, the character identity is assigned. For calculating the similarity between the input image and the template image, correlation approach is used. The Matlab function *corr2* () is used. The function compares the two images and returns a value between -1 to 1. As the value approaches to 1, it indicates that the images are more likely to be similar and as it approaches to -1, it indicates dissimilarity.

The templates are constructed by cropping the characters from image. To improve the classification accuracy, several characters are constructed for a single class character. The main reason for choosing a template matching technique is that, the characters on the plate are standardized. The characters on any car plate are the same or in others words, the characters in any plate have the same font and size. Due to this reason, template matching will minimize the work needed to classify a character to certain class. Some of the template samples used are shown in figure 3.7.

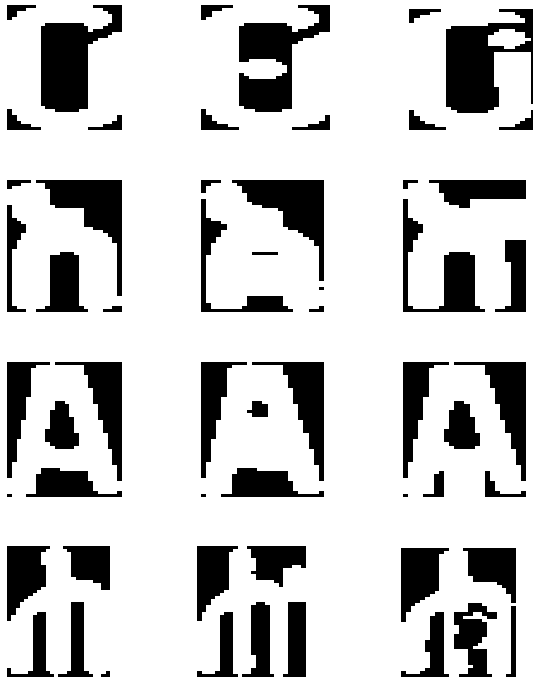


Figure 3.7: Sample template images.

After comparing the input image with the templates, class of the character is selected. The return value of the correlation process is used to classify the character, that is, return values of the comparison approaching to '1' are assumed as good match. Finally recognized characters are displayed.

Recognizing the numerals along with English alphabets is going to be a straight forward matter. However, when it comes to recognizing the Amharic characters it proved to be quite a challenge. The reason for this obstacle was that Matlab does not support all the Unicode standard international languages. The review done on this concept showed that recent versions of Matlab support all languages supported by the System Locale of the Windows OS. Unfortunately, newer versions of the windows family of Windows operating systems, Windows Vista as well as Windows 7 do not have Amharic listed in their System Locale languages while it is listed in the User Locale. This problem was a reason to search for alternative ways for displaying the recognized Amharic characters. Using a separate tool for the recognition of Amharic characters

proved to be a viable option. Java programming language is Unicode capable and displays Amharic characters. So it was decided to encode the Unicode equivalent of the recognized Amharic characters to a text file and then use a third party tool, built in Java programming language, to convert the Unicode to the proper Amharic character. However, for displaying purposes within the Matlab project, representative characters of English were used for the Amharic counterparts. At the time of recognition the Unicode equivalent of each Amharic character would be written to a text file to be later copied to the third party tool for actual conversion to Amharic letters. The tool for converting the Amharic Unicode to the Amharic character was taken from Open source code [27].

CHAPTER 4

RESULT AND DISCUSSIONS

A set of sample images are collected and used to test the developed system. The results of the test can be classified into two categories. Plate region detection and extraction result and the OCR result. The result of each stage will be discussed.

4.1. PLATE REGION DETECTION AND EXTRACTION

Static snapshots of cars were captured for testing purpose. Standing cars with few regional plates have been included. At first, the aim was to include all the regional car plates but it was difficult to obtain all of the regional plates. This is because most of the samples in Addis Ababa are not regional plates. However, few regional samples were found. In addition, some samples were also obtained from internet. With a total sample of 19 images, the result of the test can be summarized as shown in table 4.1. The accuracy percentage of the result was calculated using the formula,

$$\text{Accuracy \%} = \frac{\text{Total number of correctly detected plates}}{\text{Total number of plates}}$$

Table 4.1: Test results for plate region detection and extraction.

Total number of plates	23
Number of plate regions correctly detected and extracted	19
Number of plates that failed to be detected and extracted	4
Weighted accuracy of detected plates	82.6%

The main causes for the extraction failure are the contrast or brightness of the image and distance from camera. Images with low contrast have high chance to fail. This is due to the fact that the binary image will have more noise. Also, as the distance increases, the plate region will be small. This problem can be minimized by setting the distance value between the camera and the vehicle. Some samples for correctly extracted and which failed are shown in figure 4.1 and figure 4.2.

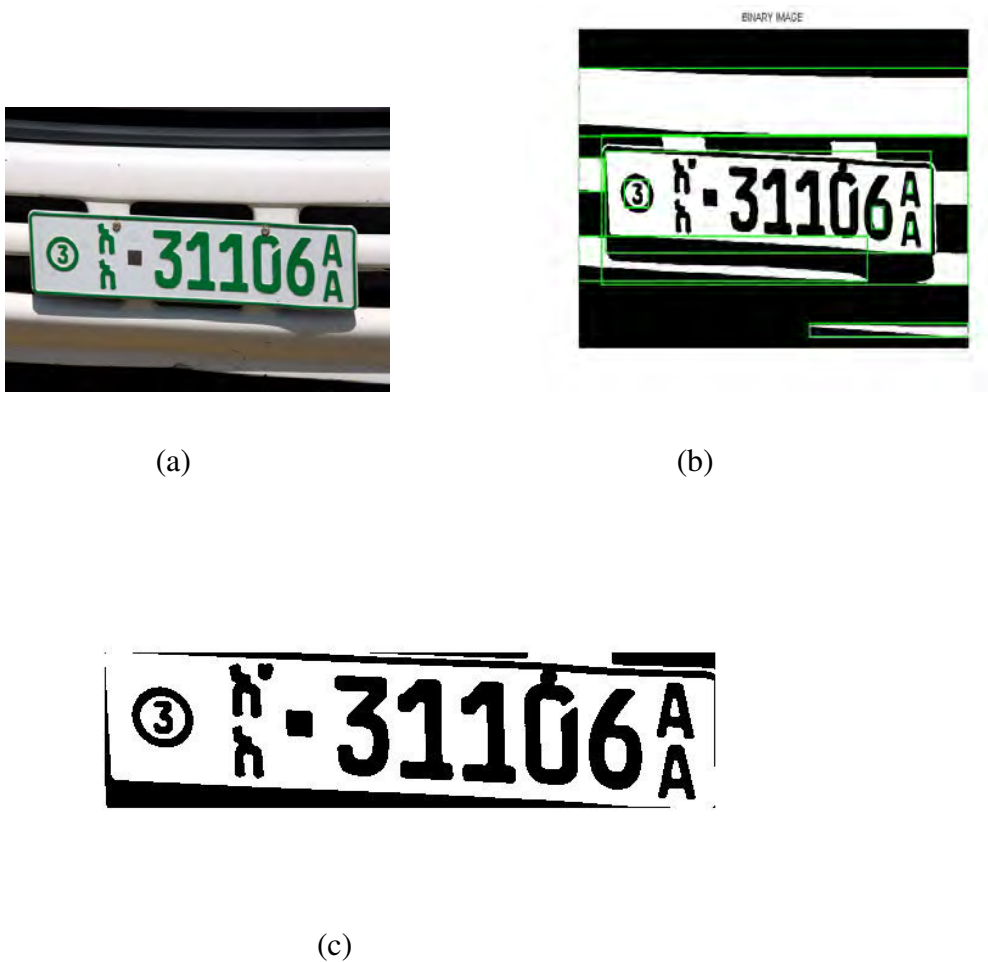


Figure 4.1: Correctly detected and extracted plate images (a) original image, (b) the binary image and (c) Plate region

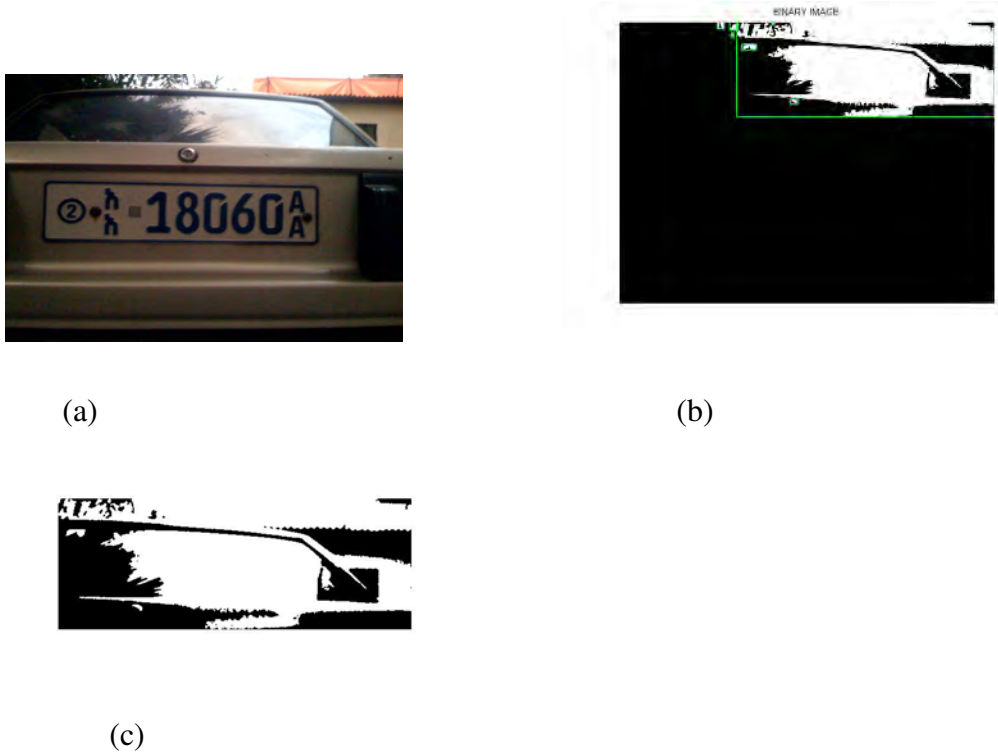


Figure 4.2: Incorrect detection and extraction (a) original image, (b) The binary image (c) Incorrect region.

4.2. OPTICAL CHARACTER RECOGNITION SYSTEM

The accuracy of the OCR in CPR system is directly influenced by the plate extraction system. This means that if the plate region fails to extract, the output of the OCR will be incorrect sample.

The sample plates, which are used for plate extraction purpose, are also used for checking the OCR recognition rate. The input for the OCR is the plate image which contains characters (numbers, Amharic and English alphabets). The result of the OCR is summarized as table 4.2. Also the overall CPR system accuracy is calculated according to table 4.3.

Table 4.2: Test result of the OCR system

Total Number of plates	23
Number of plate regions correctly detected and extracted.	19
No of characters per plate	10
Total Number of character	190
Total Number of characters correctly recognized	156
Percentage of correctly detected characters	82.11%

The overall result of the system is summarized in table 4.3.

Table 4.3: CPR Test Result

Plate Sample	Plate Region Recognized??	Total number of characters recognized	CPR Result
Sample 1	Yes	10	Success
Sample 2	Yes	10	Success
Sample 3	No	-----	-----
Sample 4	Yes	2	Failure
Sample 5	No	-----	-----
Sample 6	Yes	10	Success
Sample 7	Yes	10	Success
Sample 8	Yes	7	Failure
Sample 9	Yes	10	Success
Sample 10	Yes	10	Success
Sample 11	Yes	10	Success
Sample 12	Yes	10	Success
Sample 13	Yes	7	Failure
Sample 14	Yes	10	Success
Sample 15	Yes	9	Failure
Sample 16	No	-----	-----
Sample 17	Yes	6	Failure
Sample 18	No	-----	-----
Sample 19	Yes	10	Success
Sample 20	Yes	9	Failure
Sample 21	Yes	4	Failure
Sample 22	Yes	7	Failure
Sample 23	Yes	7	Failure
	Total Number of plate Recognized (out of 19 plates)	10	
	Overall Accuracy of CPR system	52.63%	

4.3. EVALUATION

The failure encountered in testing the CPR system can be viewed as two categories; Failure in the plate region detection and extraction and failure of OCR system.

The failure of the plate region detection and extraction is resulted due to several factors such as image quality, angle between the camera and the car, and distance between the camera and the car. The image quality directly affects the accuracy of the system. The result gives good response for high quality images. The angle between the camera and the car also influences the extraction algorithm. This is due to the variation in angle introduces variation in the illumination level and light reflection that may reduce the image quality. As a result, the extraction rate may degrade. In addition, the distance between the plate and the camera can have less influence because it can be adjusted using some techniques. But for this research, the distance used for taking the snapshots is almost constant.

The second failure category is the OCR failure. This can be further classified into three aspects, failure of the plate extractor, failure in the character segmentation process and failure in classification process. The first cause results in total failure of the OCR. Since the failure in the plate extract will transmit unwanted input to the OCR, the final output will be incorrect output.

The segmentation failure is the failure caused due to the failure of the segmentation algorithm. This failure again can be analyzed as two views, error in segmenting unwanted symbols as character and error in segmenting unwanted details in the image. The first segmentation error is caused when connected components are assumed as a single region. This error can be minimized by setting a rule for object to be segmented as a character or not. The second error is caused when unwanted details are segmented along with the characters. For example, the position of the bolts used to attach the plate on the car highly influences the segmentation result.

Failure in classification occurs when the classifier wrongly classifies/recognizes a character. Error of classifying the segmented character to the correct character class also affects the OCR performance. This happens when the characters are wrongly classified by comparing their

correlation value. For example, the Amharic letter “አ” may be misclassified as “ኢ” due to some unwanted details in the lower part of the letter “አ” .

4.4. SUMMARY

In general, from the above results, the following can be stated. The recognition rate of the CPR system is directly influenced by the quality of the input image and character recognition system. Most of the errors in the CPR are due to plate region extraction process and the errors that occur in the OCR system are the character segmentation error. From this result, it can be analyzed that the CPR accuracy is directly influenced by the plate region extraction and detection module and character segmentation and extraction module.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1. CONCLUSION

In this thesis work, CPR system for Ethiopian car plate standard is developed. The main processes in developing the systems are detecting and extracting the plate image and recognizing the characters on the plate image.

In detecting the plate region, image processing techniques are incorporated. For preprocessing phase, gray scale conversion along with median filtering technique is used. For detecting the plate region, connected component analysis (CCA) is used to analyze each region and rate the potential to be assigned as a plate region. The technique is mainly focuses on finding and segmenting white regions. Finally, using structural properties, the plate region is identified from the candidate regions.

After the plate image is extracted, it is passed to the character recognition system. For this research, OCR was used for recognizing the characters on the plate image. For segmenting the individual characters, CCA is used. After the individual characters are segmented, a template matching technique was used to classify the characters into the correct character class. Template matching techniques was preferred because the characters on the plates are standard. This property helps the template matching technique to work better.

The failures in car plate region detection and extraction are mainly caused by variation in illumination. This is because the variation in illumination results in regions brighter than the plate regions which directly results in detecting wrong regions. From this result, it can be concluded that illumination variation affects the result of CCA technique. On the other hand, the failures in the OCR are due to noise in the character such as plate bolt location, fading of the characters and others.

5.2. RECOMMENDATIONS

The following things can be enhanced to improve the performance of the CPR system.

- The plate region extraction system can be improved by including additional features which are unique to the plat region. These features may be for instance horizontal and vertical histogram of the internal region, which shows the presence of characters, can be used to identify similar regions that may affect the extraction process.
- Also improved additional features such as color based techniques can also be used.
- For the OCR, improvement in noise reduction techniques can be suggested in order to improve the segmentation and classification accuracy.

APPENDICES

APPENDIX A: GUI for Car Plate Recognition System

The interface for the Car Plate Recognition System is shown in figure A. The “Load Image” button is used to load the image for the extraction and recognition purpose. The “Extract Plate” button is used to extract the plate region from the whole image and displays the plate region. The “Recognize Plate” button is used to recognize the individual characters on the plate and displays the final out put on the lower display.

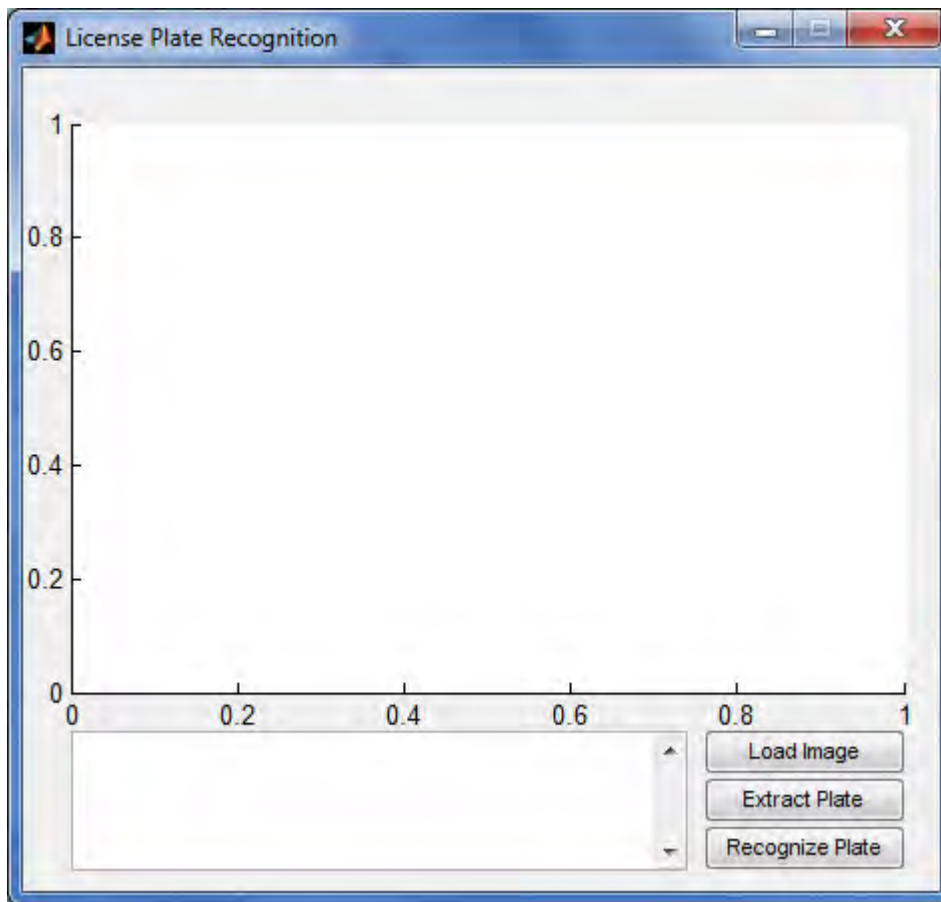


Figure A: GUI for the Car Plate Recognition system

APPENDIX B: GUI for the Amharic Unicode Converter

This interface is used to display the Amharic Character. The hex-decimal value for each Amharic character is assigned. This interface takes the hexa-decimal value and displays the corresponding Amharic character. The interface is shown in figure B. In this Example, the letters “ሐ” and “ረ” are displayed.

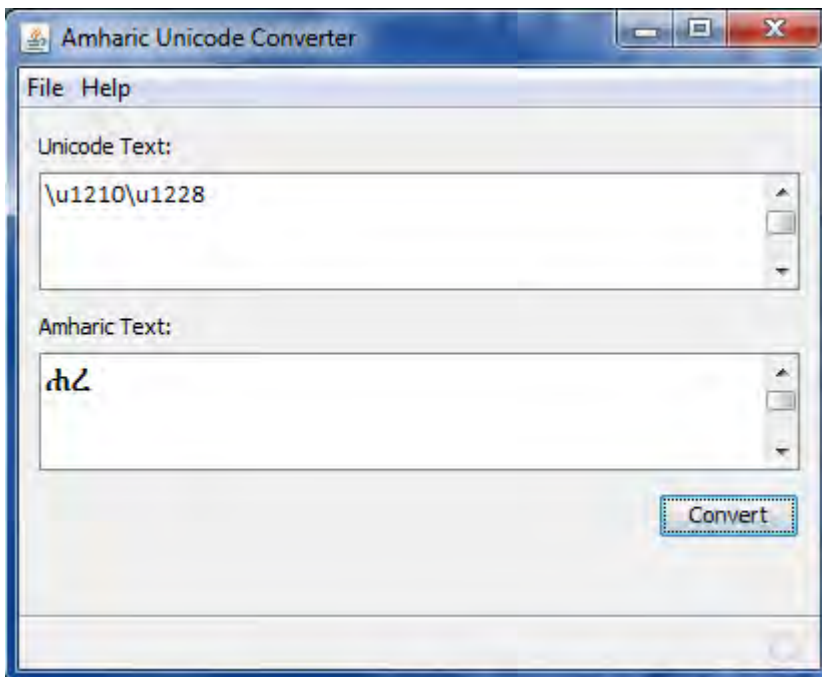


Figure B: GUI for Amharic Text Display

APPENDIX C: Sample Codes

1. This Function is used to Load the image from it location when “Load Image” button is pressed.

```
% --- Executes on button press in btnLoad.
function btnLoad_Callback(hObject, eventdata, handles)

[filename, pathname] = uigetfile({'*.jpg'; '*.bmp'; '*.gif'; '*.*'}, 'Select
License Plate Image File');
plate_image = imread([pathname, filename]);
axes(handles.image);
imshow(plate_image); %load picture
handles.plate_image = plate_image;
guidata(hObject, handles);
set(handles.txtText, 'string', '');
```

2. This Function is used to extract the plate region from the whole image when “Extract Plate” button is pressed.

```
% --- Executes on button press in btnExtract.
function btnExtract_Callback(hObject, eventdata, handles)
plate_image = handles.plate_image;
axes(handles.image);

if size(plate_image,3)==3 %RGB image
gray=rgb2gray(plate_image); %converting the RGB image to grayscale.
end

gray=medfilt2(gray, [3,3]); % Median filter to reduce noise.
[m,n]=size(gray); % calculating the size of the gray image.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Threshold segmentation
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for i=1:m
    for j=1:n
        if gray(i,j)>140
            new(i,j)=1;
        else
            new(i,j)=0;
        end
    end
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
image=bwareaopen(new,30); %Remove all object containing fewer than 30
pixels
[Label Ne]=bwlabel(image); %Label connected components
```

```

    propied=regionprops(Label,'BoundingBox'); %Measure properties of image
regions
    hold on
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %Plot Bounding Box
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    for n=1:size(propied,1)

rectangle('Position',propied(n).BoundingBox,'EdgeColor','g','LineWidth',2)
    end
    hold off
    pause (1)
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %Object Extraction
    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    for n=1:Ne
        [r,c] = find(Label==n);%finds indices and values which have the label
'n'
        ind=find(Label==n);
        [l,m]=size(ind);
        if l*m>30000 %Number of white pixels
            im=image(min(r):max(r),min(c):max(c));
            ind=find(im==0); %finds indices and values which have the value
'0' (black pixels)
            [bm,bn]=size(ind);
            if bm*bn>18000 %Number of black pixels comparison
                imwrite(im,'extracted.jpg');
                axes(handles.image);
            end
        end
    end
    end
    imshow('extracted.jpg');

```

3. This function is used to recognize the individual characters when the “Recognize Plate” button is pressed.

```

% --- Executes on button press in btnRecognize.
function btnRecognize_Callback(hObject, eventdata, handles)
    plate_image=imread('extracted.jpg');
    axes(handles.image);

    if size(plate_image,3)==3 %RGB image
        plate_image=rgb2gray(plate_image);%converting the RGB image to
grayscale.
    end

    threshold=graythresh(plate_image);
    plate_image=~im2bw(plate_image,threshold);% Convert to BW
    plate_image=bwareaopen(plate_image,30);% Remove all object containing
fewer than 30 pixels

```

```

word=[ ];
hexword=[ ];
[Label no_of_objects]=bwlabel(plate_image);% Label and count connected
components

for i=1:no_of_objects
    [r,c]=find(Label==i);
    ind=find(Label==i);
    [m,n]=size(ind);
    if m*n>200 %Number of white pixels
        extracted_letter=plate_image(min(r):max(r),min(c):max(c));%
Extract letter
        b=find(extracted_letter==0); %finds indices and values which have
the value '0' (black pixels)
        [bm,bn]=size(b);
        if bm*bn>100 %Number of black pixels comparison
            resized_letter=imresize(extracted_letter,[42 24]);% Resize
letter (same size as template)
            [letter hex]=Letters(resized_letter);%convert image to text
            word=[word letter];
            hexword=[hexword hex];
        end
    end
end

A1=cellstr(['\u1275\u12A2';'\u121B\u12A0';'\u1228\u1210';'\u12F5\u12D5'];]);
A2=cellstr(['\u12A2\u1275';'\u12A0\u121B';'\u1210\u1228';'\u12D5\u12F5'];]);
E1=cellstr(['RO';'MA';'TE';'OA';'NU';'UA';'DC';'RH';'MS';'PS';'RD';'FA';'GT'];]);
);
E2=cellstr(['OR';'AM';'ET';'AO';'UN';'AU';'CD';'HR';'SM';'SP';'DR';'AF';'TG'];]);
);
EA1=cellstr(['Amharic T Amharic È';'Amharic m Amharic A';'Amharic R Amharic
H';'Amharic d Amharic E'];]);
EA2=cellstr(['Amharic È Amharic T';'Amharic A Amharic m';'Amharic H Amharic
R';'Amharic E Amharic d'];]);

for i=1:size(E1)
    word=strrep(word,char(E1(i)),char(E2(i)));
end

for i=1:size(EA1)
    word=strrep(word,char(EA1(i)),char(EA2(i)));
end

for i=1:size(A1)
    hexword=strrep(hexword,char(A1(i)),char(A2(i)));
end

set(handles.txtText,'string',word);

```

```
file = fopen('AmharicUnicode.txt','wt');  
fprintf(file,'%s',hexword);
```

REFERENCES

- [1] Muhammad Tahir Qadri, Muhammad Asif. “Automatic Number Plate Recognition System For Vehicle Identification Using Optical Character Recognition”. International Conference on Education Technology and Computer, 2009.
- [2] Weijuan Wen, Xianglin Huang, Lifang Yang, Zhao Yang, Pengju Zhang. “The Vehicle License Plate Location Method Based-on Wavelet Transform”. International Joint Conference on Computational Sciences and Optimization, 2009.
- [3] Xianmin Wei. “Method of Blue-White Car Plate Locating”. Second IITA International Conference on Geoscience and Remote Sensing, 2010.
- [4] Xiaoping Li, Yinxiang Li, Jinghui Chen, Xiaoxing Lv, Jianqiang Xu. “Application of a New Improved Image Segmentation Algorithm in Car Plate Recognition”.
- [5] J.R. Parker, Pavol Federl. “An Approach To License Plate Recognition”. University of Calgary.
- [6] Abbas M. Al-Ghaili, Syamsiah Mashohor, Abdul Rahman Ramli and Alyani Ismail. “Car License Plate Detection Method for Malaysian Plates-Styles by Using a Web Camera”.2010
- [7] Lin Luo, Hao Sun, Weiping Zhou, Limin Luo. “An Efficient Method of License Plate Location”. The 1st International Conference on Information Science and Engineering (ICISE2009).
- [8] Christos Nikolaos E. Anagnostopoulos, Ioannis E. Anagnostopoulos, Vassili Loumos, and Eleftherios Kayafas. “A License Plate-Recognition Algorithm for Intelligent Transportation System Applications”. IEEE Transactions on Intelligent Transportation System (ITS), Vol. 7, No. 3, September 2006.

- [9] Kaushik Deb, Hyun-Uk Chae and Kang-Hyun Jo. "Vehicle License Plate Detection Method Based on Sliding Concentric Windows and Histogram". *Journal of Computers*, Vol. 4, No. 8, August 2009.
- [10] Jia LI, Mei XIE. "A Color and Texture Feature Based Approach to License Plate Location". *International Conference on Computational Intelligence and Security*, 2007.
- [11] C.Nelson Kennedy Babu, Member IEEE, Krishnan Nallaperumal, Senior Member IEEE. "A License Plate Localization using Morphology and Recognition". IEEE, 2008.
- [12] Farhad Faradji, Amir Hossein Rezaie, Majid Ziaratban. "A Morphological-based License Plate Location". *Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran. IEEE*, 2007.
- [13] Million Meshesha and C. V. Jawahar. "Optical Character Recognition of Amharic Documents". *Center for Visual Information Technology International Institute of Information Technology*.
- [14] Satadal Saha, Subhadip Basu, Mita Nasipuri and Dipak Kr. Basu. "A Hough Transform based Technique for Text Segmentation". *Journal of Computing*, Volume 2, Issue 2, February 2010, ISSN 2151-9617.
- [15] Ondrej Martinsky. "Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems". *Brno University of Technology, Faculty of Information Technology, Department of Intelligent Systems. B.Sc Thesis*, 2007.
- [16] S. Setumin, U. U. Sheikh, and S.A.R Abu-Bakar, Senior Member, IEEE Computer Vision, Video and Image Processing Research Lab (cvvip). "Character-Based Car Plate Detection and Localization". *10th International Conference on Information Science, Signal Processing and their Applications (ISSPA). IEEE*, 2010.

- [17] Yon-Ping Chen and Tien-Der Yeh. "A Method for Extraction and Recognition of Isolated License Plate Characters". (IJCSIS) International Journal of Computer Science and Information Security, Vol. 5, No.1, 2009.
- [18] Line Eivkil. "Optical Character Recognition (OCR)". December 2009.
- [19] Rafael C.Gonzalez, Richard E.Woods, Steven L.Eddins. Digital Image Processing Using Matlab. #pp 204-207. 2nd Edition.
- [20] Frank Y.Shih. Image Processing and Pattern Recognition: Fundamentals and Techniques. #pp 269-293. A John Wiley and Sons, INC., Publications, IEEE Press. 2010.
- [21] Yaregal Assabie Lake. "Optical Character Recognition of Amharic Text: An Integrated Approach". Addis Ababa University, School of Information Studies for Africa (SISA). June, 2002.
- [22] Parul Shah, Sunil Karamchandani, Taskeen Nadkar, Nikita Gulechha, Kaushik Koli, Ketan Lad. "OCR-based Chassis-Number Recognition using Artificial Neural Networks". IEEE 2009.
- [23] Majid M. Altuwaijri, Magdy A. Bayoumi. "Arabic Text Recognition Using Neural Networks". Center for Advanced Computer Studies University of Southwestern Louisiana. 1996.
- [24] Dr. John Cowell and Dr. Fiaz Hussain. "Amharic Character Recognition using a Fast Signature Based Algorithm". Proceedings of the Seventh International Conference on Information Visualization (IV'03). IEEE 2003.
- [25] Guoqiang Peter Zhang. "Neural Networks for Classification: A Survey". IEEE Transaction on Systems, Man and Cybernetics- Part C: Applications and Reviews, Vol. 30, No.4. November 2000.

[26] Alasdair McAndrew. An Introduction to Digital Image Processing with Matlab. School of Computer Science and Mathematics, Victoria University of Technology. Notes for SCM2511 Image Processing. #pp 145. 2004.

[27] www.senamirmir.org accessed on May, 2011.

[28] Prasad S.Halgaonkar. “Connected Component Analysis and Change Detection for Images”. International Journal of Computer Trends and Technology- May to June Issue 2011

DECLARATION

I, the undersigned, hereby declare that this thesis is my original work performed under the supervision of Dr. Assefa Dagne, has not been presented as a thesis for a degree program in any other university and all sources of materials used for the thesis are duly acknowledged.

Name: Huda Zuber Ahmed

Signature: _____

Place: Addis Ababa

Date of submission: November, 2011

This thesis has been submitted for examination with my approval as a university advisor.

Advisor's Name: Dr. Assefa Dagne

Signature: _____