

**UNIVERSITI TEKNOLOGI MARA**

**AUTOMATED LANE DETECTION  
OF GEL ELECTROPHORESIS  
IMAGE USING FALSE PEAK  
ELIMINATION**

**ROS SURYA BINTI TAHER**

Thesis submitted in fulfilment  
of the requirements for the degree of  
**Doctor of Philosophy**

**Faculty of Computer and Mathematical Sciences**

December 2017

## ABSTRACT

Large numbers of previous work regarding the study of lane detection in DNA gel image have been proposed and performed on good quality images. Current lane detection methods that are available do not accommodate techniques that can be performed automatically on poor DNA gel image. Lane detection is the first step in any gel image analysis techniques which involved tedious and time-consuming tasks. The accuracy of this step is often compromised by technical variation inherent to DNA gel image. For that reason, the aim of this thesis is to identify and propose a method that is effective in detecting the lane in poor DNA gel image of plants. The imperfection of DNA gel image caused by the electrophoresis or during the acquisition of the gel image causes many types of noises, which contaminate the resulting image. These errors and noises significantly affect the processing and analysis of the DNA gel image. The conducted experiment examines 184 poor DNA gel images collected from Agrobiodiversity and Environment Research Centre, Institut Penyelidikan dan Kemajuan Pertanian Malaysia (MARDI), Malaysia. The DNA gel images were produced by electrophoresis-based method using polymerase chain reaction (PCR)-based marker system. There are two highlighted aspects performed to achieve the objective of this thesis that are image enhancement and lane detection. The image enhancement of the poor DNA gel image is performed using two different approaches that are spatial and frequency filtering. The two approaches are compared and the quality of the enhanced images was accessed and evaluated using objective image quality metric that is peak signal-to-noise ratio (PSNR). For lane detection, we describe the convention of threshold value in the analysis of poor DNA gel image to eliminate false peak contained in the intensity profile obtained from the enhanced image data projection. A false peak elimination method was proposed in which a set of threshold interval was applied in the peak detection process to eliminate the false peak and retain the true peaks representing the lane's border. The output of this method is used to track the lane's border, further the individual lane is identified. Evaluation of the results from the proposed method in detecting the correct lanes was done by carrying out the analysis based on visual observation. Later, the performance of the method was evaluated empirically where the performance being assessed according to the discrepancy measures on the outcome of the lane detection process using confusion matrix. Based on the two approaches of image enhancement process, the average PSNR for the spatial domain filtering is 42.4727 dB whereas frequency domain filtering is 39.1417 dB. Therefore, spatial domain filtering become an exceptional approach for the enhancement of the poor DNA gel images. Further, the performance of the false peak elimination method proposed in lane detection and tracking process results with recall rate and accuracy of detecting the true peaks are 97.63% and 87.88% on poor DNA gel images while 99.57% and 95.96% on good DNA gel images. This finding shows that the proposed false peak elimination method, utilizing a set of optimal threshold interval, proves to be a promising lane detection method for both poor and good quality DNA gel images. Additionally, when the tasks of lane detection and tracking are implement automatically, the false peaks can be adequately eliminated and significantly ease the subsequent process of DNA gel image analysis.

## ACKNOWLEDGEMENT

All praise to Almighty Allah, the Most Gracious and Benevolent for giving me strength, endurance and an en route for completion of this dissertation.

I would like to express my profound respect to Prof. Madya Dr. Nursuriati Jamil for the great opportunity by supervising this thesis to me. I am grateful to her precious ideas, expertise, invaluable and positive help. I would also like to extend my appreciation to my co-supervisors, Dr. Sharifalillah Nordin for her valuable comments and guidance. I am also very much indebted with my co-supervisor in Agrobiodiversity and Environment Research Centre, MARDI, Pn. Umikalsum Mohamed Bahari for her expert advice and knowledge in investigating DNA gel electrophoresis image. I attribute the level of my PhD degree to their encouragement and efforts whilst allowing me the room to work in my own way.

My most sincere gratitude to my husband, Nor Faiz Ahmad Helimi and my sons, Rizzqy Ashman and Zickry Ezham for their endless loves, support and understanding. My deep and sincere thanks to my beloved parents, Taher Masri and Sainah Rahmat, and families for accomplishing things that I never thought I could do. It is to them I dedicate this thesis.

My thanks are also extended to all my friends and colleagues at the Faculty of Computer and Mathematical Sciences, UiTM Shah Alam for their great ideas and morale support. Credit should also go to Universiti Teknologi MARA (UiTM), Malaysia and Ministry Of Higher Education (MOHE) for the doctoral scholarship to pursue my PhD study.

Finally, I hope my thesis somewhere captures a remarkable contribution to the related parties; organizationally and individually.

# TABLE OF CONTENTS

	<b>Page</b>
<b>CONFIRMATION BY PANEL OF EXAMINERS</b>	ii
<b>AUTHOR'S DECLARATION</b>	iii
<b>ABSTRACT</b>	iv
<b>ACKNOWLEDGEMENT</b>	v
<b>TABLE OF CONTENTS</b>	vi
<b>LIST OF TABLES</b>	x
<b>LIST OF FIGURES</b>	xiii
<b>LIST OF ABBREVIATIONS</b>	xvii
<b>CHAPTER ONE: INTRODUCTION</b>	1
1.1 Overview	1
1.2 Research Background	2
1.3 Problem Statements	4
1.4 Research Questions	10
1.5 Research Objectives	10
1.6 Scope Of Study	11
1.7 Research Significance	12
1.8 Contribution Of Research	13
1.9 Overview Of The Thesis	16
<b>CHAPTER TWO: DNA GEL IMAGE</b>	19
2.1 DNA In Plants	19
2.2 DNA Fingerprint In Plants	20
2.2.1 DNA Fingerprinting Technique	22
2.2.1.1 Plant DNA Extraction	23
2.2.1.2 PCR Protocol	24
2.2.1.3 Restriction Fragment length Polymorphism (RFLP) Protocol	26
2.2.1.4 Random Amplified Polymorphic DNA (RAPD) Protocol	28

# CHAPTER ONE

## INTRODUCTION

### 1.1 OVERVIEW

One of the important large-scale activities that use bioinformatics is genomics that refers to the sequencing and analysis of deoxyribonucleic acid (DNA) sequences in an organism. In plants, the use of specific type of DNA sequence (genomic target region) have been utilized in the DNA gel imaging technique to reveal the specific DNA profile for particular organisms at the molecular level. The result from the DNA gel electrophoresis image or DNA gel image shows the capability to identify unique banding patterns of DNA base pair. These patterns will be useful in scoring the DNA bands particularly to estimate the genetic relatedness, to identify the plant genomics, to reveal the genetic make up of an organism and to analyse the segregation and linkage for genetic mapping.

However, the banding patterns in the DNA gel image present a great challenge to computer interpretation due to the quality and pattern of the images. Therefore, an effective analysis of the DNA gel images emerges as one of the most important applications in biological routine despite all the challenges, to reduce human error and increase the speed and accuracy of data evaluation.

The aim of the research is to identify and propose a method that is effective to detect lanes in poor plant DNA gel image. The imperfection of DNA gel image caused by the electrophoresis or during the acquisition of the gel image causes many types of noises, which contaminate the resulting image. These errors and noises can significantly affect the processing and analysis of the DNA gel image.

The research is conducted by implementing the techniques of image processing and subsequent analysis to extract quantitative information from the DNA gel image images. Further, the detection of lanes can be conducted quantitatively using method proposed in this research. The plant DNA gel images were obtained from Agrobiodiversity and Environment Research Centre, MARDI.