



**UNIVERSITI PUTRA MALAYSIA**

**DETECTION OF MICROCALCIFICATIONS USING WAVELET  
TECHNIQUES IN MAMMOGRAM DIGITAL IMAGES AND  
DEVELOPMENT OF A SOFTWARE FOR EVALUATION OF  
RADIOLOGIST FINDINGS**

**MAJDI TAYSIR HASAN AL-QDAH.**

**FK 2006 22**

**DETECTION OF MICROCALCIFICATIONS USING WAVELET  
TECHNIQUES IN MAMMOGRAM DIGITAL IMAGES AND DEVELOPMENT  
OF A SOFTWARE FOR EVALUATION OF RADIOLOGIST FINDINGS**

**By**

**MAJDI TAYSIR HASAN AL-QDAH**

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,  
in Fulfilment of the Requirement for the Degree of Doctor of Philosophy**

**January 2006**



To my family

Abstract of thesis submitted to the Senate of Universiti Putra Malaysia in fulfillment  
of the requirement for the degree of Doctor of Philosophy

**DETECTION OF MICROCALCIFICATIONS USING WAVELET TECHNIQUES IN  
MAMMOGRAM DIGITAL IMAGES AND DEVELOPMENT OF A SOFTWARE  
FOR EVALUATION OF RADIOLOGIST FINDINGS**

By

**MAJDI TAYSIR HASAN AL-QDAH**

**January 2006**

**Chairman: Associate Professor Abd Rahman Ramli, PhD**

**Faculty: Engineering**

This thesis describes a method of using wavelets in the detection comparison of breast cancer among the three main races in Malaysia: Chinese, Malays, and Indians followed by a system that records and evaluates the radiologist's findings over a period of time to gauge the radiologist's findings in confirming breast cancer cases. A comparison was carried out among few different wavelets to find out the best filter for detection for all three Malaysian races' mammograms. The detection of three filters was presented to three expert radiologists to confirm the best filter detector. As a result, the db4 wavelet was utilized to detect microcalcifications in mammogram digital images obtained from a Malaysian women sample. The wavelet filter's detection evaluation was done by visual inspection to confirm the detection results of those pixels that corresponded to microcalcifications. Detection

was counted if the wavelet detected pixels corresponded to the radiologically identified microcalcification pixels. The findings suggest that no one race mammograms are easier for wavelets' detections of microcalcifications and for the radiologist confirmation. After the radiologist's detection confirmation a new client-server radiologist recording and evaluation system is designed to evaluate the findings of the radiologist over some period of cancer detection working time. It is a system that records the findings of the Malaysian radiologist for the presence of breast cancer in Malaysian patients and provides a way of registering the progress of detecting breast cancer of the radiologist by tracking certain metric values such as the sensitivity, specificity, and Receiver Operator Curve (ROC).

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia  
sebagai memenuhi keperluan untuk ijazah Doktor Falsafah.

**PENGenalPASTIAN KALSIFIKASI MIKRO MENGGUNAKAN  
TEKNIK WAVELET DI DALAM MAMMOGRAPHY DIGITAL DAN  
PEMBANGUNAN PERISIAN UNTUK PENILAIAN PENEMUAN PAKAR  
RADIOLOGI**

Oleh

**MAJDI TAYSIR HASAN AL-QDAH**

Januari 2006

**Pengerusi : Profesor Madya Abd Rahman Ramli, PhD**

**Fakulti : Kejuruteraan**

Tesis ini menerangkan kaedah menggunakan wavelet di dalam pengenaltastian perbandingan kanser payu dara di kalangan tiga bangsa utama di Malaysia : Cina, Melayu dan India disusuli dengan sistem yang merekod dan menilai penemuan pakar radiologi dalam suatu jangkamasa tertentu yang boleh membuat penilaian dan mengesahkan kes-kes kanser payu dara. Perbandingan dibuat di antara beberapa wavelet berbeza untuk mengetahui penapis terbaik untuk pengenaltastian mammogram ketiga-tiga bangsa di Malaysia. Pengesanan tiga penapis dipersembahkan kepada tiga orang pakar radiologi untuk mengenalpasti pengesanan penapis yang terbaik. Keputusannya, wavelet db4 digunakan untuk mengesan kalsifikasi mikro dalam imej-imej digital mammogram yang diperolehi dari sampel wanita

Malaysia. Penilaian pengesanan penapis wavelet dijalankan menggunakan pemeriksaan visual untuk memastikan keputusan penilaian piksel yang berpadanan dengan kalsifikasi mikro. Pengesanan diambilkira jika piksel yang dikesan menggunakan wavelet berpadanan dengan piksel kalsifikasi mikro yang telah dikenalpasti secara logik-radiologi. Penemuan mencadangkan kemudahan untuk mengesan wavelet bagi kalsifikasi mikro dan pemastian dari pakar radiologi pada mammogram tidak dipengaruhi oleh ketiga-tiga bangsa. Selepas pemastian pengesanan dari pakar radiologi, suatu sistem baru pelanggan-pelayan yang berfungsi merekod dan menilai direkabentuk bagi membuat penilaian penemuan-penemuan radiologi dalam suatu jangkamasa kerja pengesanan kanser. Sistem ini merekod penemuan pakar radiologi di Malaysia bagi pesakit-pesakit kanser payu dara di Malaysia dan menyediakan satu cara meregister progres dalam pengesanan kanser payu dara oleh pakar radiologi dengan memerhatikan nilai metrik tertentu seperti kepekaan, keperincian dan lengkok penerima operator.

## ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Associate Professor Dr. Abd Rahman Ramli. This work would not have been completed without his supervision, support, and encouragement especially in guiding me throughout writing this thesis. Also, my special thanks go to the members of the supervisory committee Associate Professor Dr. Rozi Mahmud who provided me with valuable mammogram images, gave me her radiologist point of view, and supervised my progress throughout this work and also to Dr. Rahmita Wirza for her supervision and support.

I would like to recognize the institutions of the National Cancer Center in Kuala Lumpur, Selayang Hospital, and the Radiology Department at University Putra Malaysia that contributed expertise and images to this work; also I would like to express my gratitude to the doctors working at those facilities for contributing knowledge and consultation to this work. I would also like to express my gratitude to the staff of the graduate school for their assistance and directions in rapping up this work. My thanks go also to all of the individuals at the Computer and Communication System Department in the Engineering Faculty who have been very supportive. Finally, my appreciations go to all my friends in Malaysia who have shared me the joy of the journey.



## TABLE OF CONTENTS

	Page
<b>DEDICATION</b>	ii
<b>ABSTRACT</b>	iii
<b>ABSTRAK</b>	v
<b>ACKNOWLEDGEMENTS</b>	vii
<b>APPROVAL</b>	viii
<b>DECLARATION</b>	x
<b>LIST OF TABLES</b>	xiv
<b>LIST OF FIGURES</b>	xv
<b>LIST OF ABBREVIATIONS</b>	xix
 <b>CHAPTER</b>	
 <b>I INTRODUCTION</b>	
1.1 Introduction	1
1.4 Scope and Motivation of the Research	3
1.5 Problem Statement	4
1.2 Objectives	6
1.6 Thesis Organization	8
 <b>II LITERATURE REVIEW OF MAMMOGRAPHY PROPERTIES AND COMPUTER AIDED DIAGNOSIS (CAD) IN DIGITAL MAMMOGRAPHY</b>	
2.1 Mammography and its Properties	9
2.2 Breast Density	12
2.3 Features of Breast Abnormalities in Mammograms	13
2.3.1 Microcalcifications	14
2.3.2 Circumscribed Masses	16
2.3.3 Spiculated Lesions	18
2.4 Breast Cancer in Malaysia	19
2.5 Breast Imaging Reporting and Data System (BIRADS)	20
2.6 Mammogram Inspection by the Radiologist	20
2.7 Radiologists' Detection Evaluation Criteria	22
2.8 CAD Research in Digital Mammography	24
2.9 Applications of Wavelets	47
2.10 Programming Implementation under Matlab and JAVA	51
2.11 Summary	52

<b>III</b>	<b>THE WAVELET THEORY AND IMPLEMENTATION ARCHITECTURE</b>	
3.1	Introduction	55
3.2	Wavelet Transform-Mathematical Background	56
3.3	Implementation of Wavelets using Filter Banks	62
3.4	Wavelet Families	64
3.4.1	Daubechies Wavelets	66
3.4.2	Biorthogonal Wavelets	67
3.4.3	Coiflets Wavelets	68
3.4.4	Symlets Wavelets	69
3.5	Orthogonal Wavelet Representation versus Biorthogonal Wavelet Representation	70
3.6	Signal Resolution	70
3.7	Thresholding	71
3.8	Receiver Operator Curve (ROC)	73
3.9	Implementation Strategies Using JAVA	75
3.10	Remote Method Invocation (RMI)	79
3.11	Implementation Architecture	83
3.12	Conclusion	84
<b>IV</b>	<b>METHODOLOGY</b>	
4.1	Introduction	85
4.2	Design and Material Implementations	88
4.3	Use Cases of the system	89
4.4	Detection module	91
4.4.1	Detection Methods and Approaches	91
4.4.2	Wavelet Filters and Race Sampling	100
4.4.3	Designing the Detection Graphical User Interface	102
4.5	Evaluation module	104
4.5.1	Evaluation System Metrics	104
4.5.2	Receiver Operator Curve (ROC)	109
4.5.3	Evaluation System Rules	114
4.6	Evaluation System Design Flow-Charts	120
4.6.1	Entity-Relationship Diagram (ERD)	120
4.6.2	Data Flow Diagrams (DFD)	123
4.6.3	Server Flow Chart	132
4.6.4	Client Flow Chart	133
4.6.5	Hierarchy of Modules	136
4.7	Summary	139
<b>V</b>	<b>RESULTS AND DISCUSSIONS</b>	
5.1	Detection Results and Discussions	140
5.2	Evaluation Results and Discussions	160

5.3 Limitations of the Study	175
5.4 Brief Summary	176
<b>VI CONCLUSION</b>	
6.1 Concluding Remarks	177
6.2 Suggestions for Future Research	180
<b>REFERENCES</b>	181
<b>APPENDICES</b>	189
<b>BIODATA OF THE AUTHOR</b>	195

## LIST OF TABLES

Table	Page
3.1 Data for evaluation	75
3.2 Import RMI Classes	81
4.1 Patients' Data Table for Evaluation	111
4.2 Calculation process for TPR and FPR at each BIRAD	112
4.3 Summary of the Calculated Results of TPR and FPR	113
4.6 Calculation Results of TPR and FPR	127
5.1 The Scoring Range Used by Radiologists	145
5.2 The Averages of Detection Given by Three Radiologists	146
5.3 The Mammogram Sample From the Three Races	153
5.4 The Statistical Values of Detection for Each Race	155
5.5 The Sensitivity and Specificity for Each Race Sample	155
5.6 The AUC Value for the ROC's for Each Race	158
5.7 Patient's Mammogram Examination	165
5.8 Sample Patient's Diagnosis Data	170
5.9 Classification of Patients Under Different Age Groups	175

## LIST OF FIGURES

Figure	Page
2.1 Normal Mammogram Cutout	13
2.2 Unprocessed Mammogram Image Cutout with Calcifications	14
2.3 Two Basic Types of Malignant Microcalcifications	15
2.4 Benign Microcalcifications.	15
2.5 Malignant Masses	17
2.6 Benign Masses	17
2.7 Malignant Spiculated Lesions	18
2.8 Benign Spiculated Lesions	19
2.9 A Cleaned signal After Eliminating Some Noise	48
2.10 A Wavelet Decomposition of an Image	50
3.1 A Sine Wave and a Daubechies 2 wavelet (db2)	60
3.2 Schematic Area Tiling of the Time-Frequency Plane	61
3.3 A Signal with Discontinuity at $t=0.5$	61
3.4 The Fourier and Wavelet Representation of a Signal	62
3.5 Decomposition of a Signal Using a Filter Bank	63
3.6 Multi-level Signal Decomposition of a Signal	64
3.7 Reconstruction of a Signal Using a Filter Bank	64
3.8 A Relationship between Scale and Shape of the Wavelet	66
3.9 Nine Different Wavelets of the Daubechies Family	67

3.10	The Different Biorthogonal Family of Wavelets	68
3.11	The Five Different Wavelets of the Coiflet Family	69
3.12	The Seven Different Wavelets of the Symlet Family	70
3.13	JDBC-to-Database Communication Paths	78
3.14	A General RMI Architecture	82
3.15	Three-Tier Application Architecture	83
4.1	Block Diagram of Detection and Evaluation System	89
4.2	The Use Case Diagram of the System	91
4.3	The Scheme for Wavelet Based Image Processing	93
4.4	Detection Flow Chart	94
4.5	The Subbands After Three Levels of Wavelet Decomposition	98
4.6	Decomposition of WT without Downsampling	99
4.7	Reconstruction of WT without Upsampling	99
4.8	GUI Development Environment in Matlab	103
4.9	Running the GUI under Matlab Environment	103
4.10	ROC Calculation and Derivation Procedure	110
4.11	ROC curves	113
4.12	Interaction between the Diagnosis and the Examination Entity	115
4.13	Business Rule Defining the Diagnosis Done by each Radiologist	115
4.14	Every Diagnosis Belongs to One of Five Assessments	116
4.15	Every Diagnosis Has One and Only One of Ten Findings	117
4.16	One or Several Diagnosis Must be Associated with a Patient	117

4.17	Every Patient Belongs to Only One Race	118
4.18	Each Female Patient Has One of Five Types of Breast Tissue	119
4.19	Each Patient Can Have One or Many Mammogram Examinations	119
4.20	ERD of the evaluation module	121
4.21	Context Diagram of the Evaluation Module	122
4.22	Level 0 Data Flow Diagram of the Evaluation Module	123
4.23	User Login Data Flow Diagram	124
4.24	Level 1 Data Flow Diagram of Viewing Diagnosis	125
4.25	Level 1 DFD of Editing Diagnosis	126
4.26	Level 1 DFD of Adding New Mammogram Examination	127
4.27	Level 1 DFD of Statistics Calculations	128
4.28	Level 1 DFD of List of Radiologists	128
4.29	Level 1 DFD of Maintaining Patients	129
4.30	Level 1 DFD of Maintaining Radiologists	130
4.31	Level 1 DFD of maintaining Authorized Users Login Accounts	131
4.32	Server Flow Chart	133
4.33	Client Flow Administration Chart	134
4.34	Client Flow Radiologist Chart	135
4.35	Client Side Module Hierarchy	137
4.36	Server Side Module	138
4.37	The RMI module	139
5.1	A Mammogram Image with Calcifications Visible	147

5.2	The Detection of the db4 Wavelet Filter	147
5.3	The Detection of the Coif2 Filter	148
5.4	The Detection of the Sym4 Filter	148
5.5	A Mammogram Image With Calcifications Visible	149
5.6	The Detection of the db4 Wavelet Filter	149
5.7	The Detection of the Coif2 Filter	149
5.8	The Detection of the Sym4 Filter	150
5.9	The db4 Mother Wavelet	152
5.10	ROC Curve for the Malay Race	156
5.11	ROC Curve for the Chinese Race	156
5.12	ROC Curve for the Indian Race	157
5.13	Successful Connection to the Server	163
5.14	Main Interface for a Radiologist	165
5.15	Mammogram Examinations' List	167
5.16	Patient's Search Under Some Constraints	168
5.17	One Diagnosis Finalization Screen	169
5.18	All Diagnosis List	171
5.19	ROC curves	173



## LIST OF ABBREVIATIONS

RMI	Remote Method Invocation
NCR	National Cancer Center
BIRADS	Breast Imaging Reporting and Data System
ACR	American College of Radiology
ROC	Receiver Operator Curve
ROI	Region of Interest
ATMTN	Asynchronous Transfer Mode Telemammography Network
CAD/DSP	Diagnosis-Detection/Digital Signal Processing
DICOM	Digital Imaging and Communications in Medicine
GUI	Graphical User Interface
MRRES	Malaysian Radiologist Records and Evaluation System

## CHAPTER I

### INTRODUCTION

#### 1.1 Introduction

Breast cancer is one of the most common cancers for women in different countries and Malaysia is no exception. Different women across the world have variations in breast cancer tissue depending on the body size, diet, reproductive characteristics such as age of menarche, number of children, and age of menopause. Also, existence rates vary from one race to another; higher rates of breast cancer have been reported in white women more than in black women [1]. A study done by University Malaya Medical Center in Kuala Lumpur (HUKL) in the year 2000 has shown that 60% of 952 cancer patients admitted to the UHKL in the years 1993 to 2000 were Chinese patients. It was concluded that the incidence of breast cancer in Chinese appears to be higher than the other two races, namely Malays and Indians [1]. It was also shown that 30-40% of all patients were in the late stages of the cancer. After comparing the data with Caucasians' data in the United States and Europe it was suggested that Malaysian women tend to get the disease at an early age with smaller tumor sizes. Also the findings suggest that breast cancer in Malaysians is more aggressive type than Caucasians. Furthermore, it can be said that women self-examination has sometimes shown shortcomings in that it allowed the cancer to grow to dangerous points before detection is made especially in

Malaysian women. This makes it clear that a method that makes an earlier detection of breast cancer possible is always needed. Nowadays this method is through screening with mammography.

Regarding computer methods in mammography, it is a difficult task to design methods of automatic computer detection because of the nature of mammograms. The mammograms are images of high resolution and low contrast with a great variation in the grayscales of different mammograms. When designing a computer system for analysis of mammograms, it is necessary to find methods suitable for locating early signs of breast cancer, microcalcifications. In a mammogram, radiologists try to identify calcifications as indicators of early cancer. In some mammograms the calcifications are seen as white spots on a dark gray background while in other mammograms are visible as brighter gray spots on a slightly darker gray background. There may also be other bright regions not associated with calcifications, which makes straightforward methods applied to other medical images inappropriate for detection of early signs of cancer, microcalcifications, and there is a need to help radiologists obtain better diagnosis by providing special computer aided tools. One good tool for analysis of medical images is wavelets' analysis [2]. The wavelet transform is a useful mathematical tool that currently has been applied in different applications of image processing including the detection of cancers in mammograms.

## 1.2 Scope and Motivation of the Research

Since it has been reported by the Malaysian National Cancer Center (NCR) [1] that the incidence of breast cancer is lower in Malay women compared to Chinese and Indian women, the race factor in mammography has to be investigated; especially if the risk of cancer in Chinese women is related to the race factor or it is related to the lifestyle, reproductive practices, age, diet, genetics, which are factors that have been shown to influence the occurrence of cancer not only in the breast but also all over the body and among different races of the world. The question might be asked whether it is also difficult to diagnose breast cancer in one race of the Malaysian society more than other races. This work will concentrate on the study and comparison of the three Malaysian race's mammograms in relation to the detection of breast cancer. It uses the wavelet detection techniques to compare the easiness or difficulty of detecting early signs of cancer in Malaysian mammogram images. The work is not concerned with the study of pattern classification of the clustered shapes of calcifications within a mammogram image but more on the detection comparison of the wavelet filter for different Malaysian races' mammograms. Also, since the definite results of studying any factor related to cancer can only be obtained after a long period of time, real data by working radiologists has to be obtained and evaluated related to the race factor study or even related to other factors' studies that can be done by other researchers working in the field. Therefore, this work designs and implements a

separate evaluation system that allows radiologists to record and track their findings for a period of working time with system automatic evaluation at each period of the radiologist findings. The design of such system can act as a computer aid tool for radiologists and medical institutions since medical computer specialists currently are trying to develop complete medical information systems that all doctors being an orthopedics, pediatric, or a radiologist and other medical staff can use within a medical environment. This work will help design the radiologist component of the information system for use of medical institutions and for researchers' utilization to study certain factors' relation to breast cancer.

### **1.3 Problem Statement**

Microcalcifications form the early signs of cancer and they appear as bright small white spots in a mammogram image; but they are not easily detected by radiologists which results in many missed diagnoses. Several techniques and methods have been proposed for enhancing and extracting microcalcification from mammogram images including neural networks and fuzzy techniques (Nagel et al. [40], Cheng et al. [39], and Verma and Zakos [72]); morphological operations (Betal [37], Pohlman et al. [55], Quadrates & Sacritan [34], and Smith et. al [35]); fractal methods (Sang & Jun [51] and Huai et. al [50]); and lastly wavelet based techniques (Liu et al. [61], Laine et al. [62], Chen et al. [63], Strickland and Hahn [64], Yoshida

et al. [65], Murat. et. al [67], and Werapon & Kosin [66]). Wavelets specifically have been used in many applications and in the use of identifying small features within a mammogram image as they can perform better in the detection of those fine features than other conventional methods such as unsharp masking, morphological operations, matched filters, and multi scale segmentation Murat et al. [67]. Wavelets have good filtering abilities since they use multiresolution property to analyze images that allows them to zoom on in details within an image at different subbands. But more work is needed to investigate the suitable wavelet filter(s) for the analysis and its performance by comparing wavelets from the same family and among different families; also more work is needed to investigate the breast tissue characteristics' variations among different races in cancer detection because of the intensity contrast values and their effects on the detection capabilities of the filters. Moreover, there is no method that can accomplish a 100% detection and positive presence of breast cancer especially if the microcalcifications are not clustered; therefore, the radiologist should be involved in the decision-making and the final results of identifying the presence of benign or malignant cancers in any detection system even though it is understood that malignant tumors could arise if the microcalcifications are usually clustered it is not necessarily true for all clustered cases; furthermore, the cluster size and shape is not clearly defined and could vary from few pixels to more number of pixels in order