DEDICATION

I dedicate my thesis to my family and many friends.

A special feeling of gratitude to my loving parents: Bushara Garma and Hikmet Mohamed Ahmed whose words of encouragement.

My sister Fawzia, have never left my side, my aunt Omalhasun Mohamed Ahmed and my uncle Bakry Mohamed.

I also dedicate this dissertation to my many friends who have supported me throughout the process: Maha Almoona, Alaa Ali, Arwa AlaaEldeen, Marwa AbdoElmonim and Sabreen Mobark.

I dedicate this work and give special thanks to my Teachers.

And to all women suffering from breast cancer
ACKNOWLEDGEMENTS

I would like to give thanks to Allah for guiding my directions throughout this thesis, and for granting me more knowledge, more insight and more enlightenment.

I am deeply indebted to my supervisor Dr. Mawia Ahmed Hassun for his wisdom to move me forward and help me to reach this level.

Special thanks to Dr. Mohamed Al fadel, his countless hours of reflecting, reading, encouraging, and most of all patience throughout the entire process.

I would like to acknowledge and thank Dr. Mohamed Yagoub, Dr. Elyas Sidig and Dr. Alaa Eldeen Awooda for dedicating to me some of their time and experience.

Special thanks goes to the members of staff Biomedical Engineering department for their continued support.

Finally I would like to thank the engineer Alla Ali Alnasry that assisted me with this thesis.
ABSTRACT

The breast cancer is a serious public health problem among women in the world. Mammogram breast X-ray is considered the low cost and most reliable method in early detection of breast cancer. In this thesis an approach is proposed to develop a Computer-Aided Diagnosis (CAD) system that can be very helpful for radiologist in diagnosing.

This work has tried to analyze the texture of mammography images taken from Mini MIAS data base and to find the values of various parameters of texture. Two features types, Haralick’s features based on spatial grey level dependency (SGLD) matrix and based wavelet coefficients are applied for classification of each Regions of Interest (ROIs).

The proposed method for detection of breast cancer on digital mammogram classified the normal breast tissues into three classes which: fat, glandular and dense and then into normal and abnormal classes. The features discriminating to detect abnormal from normal tissues was determined by stepwise linear discriminant analysis classifier (LDA).

This study investigates whether the texture could be used to discriminate among the various breast tissue types. The proposed method focuses on SGLD matrix as parameters for texture analysis which achieved the highest accuracy that 95.7% for classification of breast tissues on digital mammograms. This is an important step in the development of a CAD for mammograms analysis being developed.
المستخلص

يعتبر مرض سرطان الثدي من أخطر مشاكل الصحة العامة التي تصيب النساء في العالم. ففحص الماموغرافي هو عبارة عن فحص بالأشعة السينية ويستخدم لفحص الثدي. إن الفحص المبكر لسرطانات الثدي التي يتم تشخيصها عن طريق فحص الماموغرافيك يزيد فرصة نجاح العلاج من هذا المرض. كما يعتبر تصوير الثدي بالأشعة السينية هو الطريقة الأقل تكلفة والأكثر موثوقية في الكشف المبكر عن سرطان الثدي. في هذه الأطروحة أقترحت طريقة لكشف سرطان الثدي عن طريق ما يسمى بنظام التشخيص بمساعدة جهاز الحاسوب (CAD system) و الذي يمكن أن تكون بمثابة جرعة فعالة للتشخيص.

وقد أعتمد في هذا العمل على التحليل اللفكي (Texture analysis) لصور الماموغرام والتي تم الحصول عليها من جمعية تحليل صور الماموغرام MIAS. استخدمت معادلات هراليك Spatial Grey Level المشتقة من طريقة إحصائية تسمى ب (Haralick Features) و طريقة تحويلية تسمى ب (Wavelet Coefficients) حيث تمكن أهميتها في التمييز بين الخلايا المريضة والسليمة والتي تحدد بواسطة معادلة التحليل الخطي (Linear Discriminant Analysis)

الطريقة المقترحة للكشف عن سرطان الثدي على الماموجرام الرقمي هي تصنيف أنفسة الثدي السليمة إلى ثلاث فئات منها: الدهون، العدد وو الأسنان الكثيفة و الأسنان المصابة بالسرطان (Haralick) وأظهرت النتائج أن أفضل طريقة في كشف سرطان الثدي هي معادلات هراليك و (Spatial Grey Level Dependency matrix) المشتقة من طريقة إحصائية تسمى ب (MIAS) التي حققت أعلى دقة قيمتها 95.7% مقارنة بالتشخيص المعمد من MIAS. وتعتبر هذه الدراسة خطوة مهمة في تطوير نظام التشخيص باستخدام الكمبيوتر لتطوير تحليل صور الماموغرام.
Table of Contents

CHAPTER ONE................................................................. Error! Bookmark not defined.
1. INTRODUCTION ......................................................... Error! Bookmark not defined.
1.1 General Overview ..................................................... Error! Bookmark not defined.
1.2 The Problem Statement ............................................... Error! Bookmark not defined.
1.3 Thesis Objectives ...................................................... Error! Bookmark not defined.
1.4 Thesis Organization .................................................... Error! Bookmark not defined.
CHAPTER TWO ................................................................. Error! Bookmark not defined.
2. THEORETICAL BACKGROUND ....................................... Error! Bookmark not defined.
2.1 Breast Anatomy ......................................................... Error! Bookmark not defined.
2.2 Breast Cancer ........................................................... Error! Bookmark not defined.
2.2.1 Breast cancer lesions ............................................... Error! Bookmark not defined.
2.2.2 Types of Breast Cancer ............................................. Error! Bookmark not defined.
2.3 Mammography .......................................................... Error! Bookmark not defined.
2.3.1 Principles of mammography ..................................... Error! Bookmark not defined.
2.3.2 Mammography equipment ........................................ Error! Bookmark not defined.
2.3.3 Full Field Digital Mammography ............................... Error! Bookmark not defined.
2.3.4 Mammogram formation .......................................... Error! Bookmark not defined.
2.3.5 Mammogram projections ......................................... Error! Bookmark not defined.
2.4 Computer Aided Detection .......................................... Error! Bookmark not defined.
2.5 Texture analysis ........................................................ Error! Bookmark not defined.
2.5.1 Structural approaches ........................................... Error! Bookmark not defined.
2.5.2 Statistical approaches ............................................ Error! Bookmark not defined.
2.5.3 Model based method .............................................. Error! Bookmark not defined.
2.5.4 Transform methods ................................................ Error! Bookmark not defined.
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6</td>
<td>Wavelet analysis</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Fourier transform (FT)</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Short-Time Fourier analysis (STFT)</td>
</tr>
<tr>
<td>2.6.3</td>
<td>Wavelet Transform</td>
</tr>
<tr>
<td>2.7</td>
<td>Multiresolution analysis</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Scale function</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Wavelet function</td>
</tr>
<tr>
<td>2.8</td>
<td>The Continuous Wavelet Transform</td>
</tr>
<tr>
<td>2.9</td>
<td>The Discrete Wavelet Transform (DWT)</td>
</tr>
<tr>
<td>2.10</td>
<td>Wavelet transform in two dimension</td>
</tr>
<tr>
<td>2.11</td>
<td>Wavelet Families</td>
</tr>
<tr>
<td>2.11.1</td>
<td>Haar</td>
</tr>
<tr>
<td>2.11.2</td>
<td>Daubechies</td>
</tr>
<tr>
<td>2.11.3</td>
<td>Biorthogonal</td>
</tr>
<tr>
<td>2.11.4</td>
<td>Coiflets</td>
</tr>
<tr>
<td>2.11.5</td>
<td>Symlets</td>
</tr>
<tr>
<td>2.11.6</td>
<td>Morlet</td>
</tr>
<tr>
<td>2.11.7</td>
<td>Mexican Hat</td>
</tr>
<tr>
<td>2.11.8</td>
<td>Meyer</td>
</tr>
<tr>
<td>2.12</td>
<td>Texture analysis</td>
</tr>
<tr>
<td>2.13</td>
<td>The Linear discriminant function</td>
</tr>
</tbody>
</table>

CHAPTER THREE

3. LITERATURE REVIEW

3.1 Statistical methods

3.2 Wavelet decomposition methods

CHAPTER FOUR

4. METHODOLOGY

4.1 Preprocessing stage

4.1.1 Mammogram Data

4.1.2 Selection of ROI

4.2 Feature Extractor

2.1.1 The Spatial Gray Level Dependency matrix
2.1.2 Wavelet Generator ............................................ Error! Bookmark not defined.
4.3 Texture Analysis .................................................. Error! Bookmark not defined.
4.3.1 The texture features from the SGLD matrix. Error! Bookmark not defined.
4.3.2 The texture features from the wavelet map... Error! Bookmark not defined.
4.3.3 The texture features from the combination step ........Error! Bookmark not defined.
4.4 Classification stage ............................................. Error! Bookmark not defined.
4.5 Proposed Algorithm............................................. Error! Bookmark not defined.
CHAPTER FIVE .......................................................... Error! Bookmark not defined.
5. RESULTS AND DISCUSSION .................................... Error! Bookmark not defined.
5.1 Results Overview .................................................. Error! Bookmark not defined.
5.1.1 Spatial gray level dependency matrix .......... Error! Bookmark not defined.
5.1.2 Wavelet Basis Selection................................. Error! Bookmark not defined.
5.1.3 Full features analysis........................................ Error! Bookmark not defined.
5.2 Discussions ......................................................... Error! Bookmark not defined.
CHAPTER SIX .......................................................... Error! Bookmark not defined.
6. CONCLUSION AND RECOMMENDATIONS .... Error! Bookmark not defined.
References................................................................. Error! Bookmark not defined.
List of Figures

Figure 2.1: The anatomy of the breast [7]
Figure 2.2: Type of microcalcifications commonly seen on mammogram[11]
Figure 2.3: Morphologic spectrum of mammographic masses [17]
Figure 2.4: Mass examples with different shapes and borders [16]
Figure 2.5: Schematic diagram of a dedicated mammography machine [21]
Figure 2.6: Film-screen receptor [21]
Figure 2.7: (a) Small-format detector system for biopsy imaging, (b) full-breast detector incorporating 12 detector modules, (c) slot detector for a full-breast scanning digital mammography system
Figure 2.8: Schematic representation of digital mammography system [21]
Figure 2.9: Two distinct mammography projections: (a) crano-caudal view and (b) mediolateral oblique view
Figure 2.10: A schematic diagram for CAD system
Figure 2.11: The Fourier Transform
Figure 2.12: The Short-Time Fourier Transform
Figure 2.13: Contrast with the (a) time-based, (b) frequency-based, (c) Short-Time Fourier transform (STFT) and (d) wavelet views of a signal
Figure 2.14: Multiresolution of a Signal: (a) An arbitrary signal can be decomposed into (b) approximation and (c) detail components
Figure 2.15: Multiresolution Hierarchical Model. A0 is the original image. A represents the approximation maps, D represents the detail maps, and the digits represent the hierarchy level
Figure 2.16: Schematic diagram of 2D wavelet transform
Figure 2.17: Image Decomposition with Wavelet Transform: (a) Decomposition level 1 and (b) Decomposition level 2
Figure 2.18: The Harrwavelet
Figure 2.19: Daubechies wavelet db2 to db10
Figure 2.20: The Coiflets wavelet: (a) coif1, (b) coif2, (c) coif3, (d) coif4 and (e)coif5
Figure 2.21: The Morlet Wavelet
Figure 2.22: The Mexican Hat Wavelet
Figure 2.23: The Meyer Wavelet
Figure 4.1: Block diagram of proposed method for detection of breast cancer
Figure 4.2: ROIs: (a) Fat, (b) Glandular and Connective, (c) Dense and (d) ROI with abnormal tissues.

Figure 4.3: Construction of the SGLD matrix for $d = 1$: (a) the original image 4×4 pixels begin by having each of its neighboring pairs examined. Part (b) the final result of the horizontal SGLD matrix $d=1$.

Figure 4.4: The propose algorithm for detection breast cancer using SGLD matrix.

Figure 5.1: A scatter plot of the four classes (Fat, Glandular, Dense, and Cancer) extracted from all of the features using the SGLD matrix.

Figure 5.2: A scatter plot of the four classes (Fat, Glandular, Dense, and Cancer) extracted from all of the features using the SGLD matrix at the same time with wavelet coefficients.

Figure 5.3: A scatter plot of the four classes (Fat, Glandular, Dense, and Cancer) extracted from all of the features using the combination between SGLD matrix and db4 wavelet at level 1.

Figure 5.4: A scatter plot of the four classes (Fat, Glandular, Dense, and Cancer) extracted from all of the features using the combination between SGLD matrix and db4 wavelet at level 2.

Figure 5.5: A scatter plot of the four classes (Fat, Glandular, Dense, and Cancer) extracted from all of the features using the combination between SGLD matrix and db4 wavelet at level 3.

Figure 5.6: Entropy versus classes using SGLD matrix. The entire data set of 300 normal and 114 abnormal ROIs is used as input cases.

Figure 5.7: Energy versus classes using SGLD matrix. The entire data set of 300 normal and 114 abnormal ROIs is used as input cases.

Figure 5.8: Inertia versus classes using SGLD matrix. The entire data set of 300 normal and 114 abnormal ROIs is used as input cases.

Figure 5.9: Inverse Different Moment versus classes using SGLD matrix. The entire data set of 300 normal and 114 abnormal ROIs is used as input cases.

Figure 5.10: Correlation versus classes using SGLD matrix. The entire data set of 300 normal and 114 abnormal ROIs is used as input cases.

Figure 5.11: Variance versus classes using SGLD matrix. The entire data set of 300 normal and 114 abnormal ROIs is used as input cases.
List of Tables

Table 5-1: Classification Results of spatial gray level dependency matrix........Error!
Table 5-2 Classification Results of db1 in horizontal direction level 2..............Error!
Table 5-3: Classification Results of db4 in horizontal direction level 1.............Error!
Table 5-4: Classification Results of db4 in vertical direction level 1Error! Bookmark not defined.
Table 5-5: Classification Results of db4 in diagonal direction level 1..............Error!
Table 5-6: Classification Results of db4 in vertical direction level 2Error! Bookmark not defined.
Table 5-7: Classification Results of db4 in diagonal direction level 2..............Error!
Table 5-8: Classification Results of db8 in horizontal direction level 1.............Error!
Table 5-9 Classification Results of db8 in diagonal direction level 1..............Error!
Table 5-10: Classification Results of db8 in horizontal direction level 2............Error!
Table 5-11: Classification Results of db8 in diagonal direction level 2..............Error!
Table 5-12: Classification Results of db8 in diagonal direction level 3..............Error!
Table 5-13: Classification Results of Harr in horizontal direction level 1...........Error!
Table 5-14: Classification Results of Harr in diagonal direction level 1.............Error!
Book mark not defined.
Table 5-15: Classification Results of Harr in horizontal direction level 3 ..........Error!

Table 5-16: Classification Results of Harr in vertical direction level 3 ..........Error!

Table 5-17: Summary of Classification Results: Normal Classification Rates for Each Feature Type. ................................................Error! Bookmark not defined.

Table 5-18: Classification Results of SGLD matrix in horizontal direction. .........Error!

Table 5-19: Classification Results of db4 in horizontal direction level 1..........Error!

Table 5-20: Classification Results of db4 in vertical direction level 1 ..........Error!

Table 5-21: Classification Results of db4 in diagonal direction level 1 ..........Error!

Table 5-22: Classification Results of db4 in horizontal direction level 2.........Error!

Table 5-23: Classification Results of db4 in vertical direction level 2.........Error!

Table 5-24: Classification Results of db4 in horizontal direction level 3 ..........Error!

Table 5-25: Classification Results of db4 in vertical direction level 3 ..........Error!

Table 5-26: Classification Results of db4 in diagonal direction level 3 ..........Error!

Table 5-27: Summary of Classification Results for each feature type. ............Error!

Table 5-28: Classification Results of the SGLD matrix combinations with db4 in horizontal vector level 1 which classified into each of the four classes. ..........Error!

Table 5-29: Classification Results of the SGLD matrix combinations with db4 in horizontal vector level 2 which classified into each of the four classes. ..........Error!

Table 5-30: Classification Results of the SGLD matrix combinations with db4 in horizontal vector level 3 which classified into each of the four classes. ..........Error!

Table 5-31: Classification Results between db4 in horizontal vector (level 1 with level 2) which classified into each of the four classes..........Error! Bookmark not defined.

Table 5-32: Classification Results between db4 in horizontal vector (level 1 with level 3) which classified into each of the four classes.........Error! Bookmark not defined.

Table 5-33: Classification Results between db4 in vertical vector (level 1 with level 2) which classified into each of the four classes. ..............Error! Bookmark not defined.

Table 5-34: Classification Results between db4 in vertical vector (level 1 with level 3) which classified into each of the four classes. ..............Error! Bookmark not defined.
Table 5-35: Classification Results between db4 in vertical vector (level 2 with level 3) which classified into each of the four classes. Error! Bookmark not defined.
Table 5-36: Classification Results between db4 in diagonal vector (level 1 with level 2) which classified into each of the four classes. Error! Bookmark not defined.
Table 5-37: Summary of Classification Results. Normal versus. Abnormal classification rates for each combination feature type. Error! Bookmark not defined.
Table 5-38: Classification Function Coefficients generated by linear discriminant functions classifier from SGLD matrix. Error! Bookmark not defined.