

**FRACTAL BASED TECHNIQUES FOR
CLASSIFICATION OF MAMMOGRAMS
AND
IDENTIFICATION OF MICROCALCIFICATIONS**

Thesis submitted to

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

in partial fulfillment of the requirements for the award of the degree of

DOCTOR OF PHILOSOPHY

in the Faculty of Technology

DEEPA SANKAR

under the guidance of

Prof. TESSAMMA THOMAS



**DEPARTMENT OF ELECTRONICS
FACULTY OF TECHNOLOGY
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY,
KOCHI-682022. KERALA, INDIA**

August 2011

Fractal based techniques for classification of mammograms and identification of microcalcifications

Ph.D Thesis in the field of Image Processing

Author

Deepa Sankar

Department of Electronics

Cochin University of Science and Technology

Kochi-682022

Kerala

India

email: deepasankar@cusat.ac.in

Supervising guide

Prof. Tessamma Thomas

Department of Electronics

Cochin University of Science and Technology

Kochi-682022.

Kerala.

India.

email: tess@cusat.ac.in

August 2011

To my Parents.....

Department of Electronics
Cochin University of Science and Technology,
Kochi-682022. Kerala, India.

Certificate

*Certified that this thesis entitled "Fractal based Techniques for Classification of Mammograms and Identification of Microcalcifications " is a bonafide record of the research work done by **Deepa Sankar** under my supervision in the Audio and Image Research Lab, Department of Electronics, Cochin University of Science and Technology, Kerala, India. The results presented in this thesis or parts of it have not been presented for the award of any other degree.*

Kochi-22
16 August 2011

Prof. Tessamma Thomas
(Supervising Guide)
Department of Electronics
Cochin University of Science and Technology
Kochi-682022

DECLARATION

*I hereby declare that the work presented in this thesis entitled "**Fractal based Techniques for Classification of Mammograms and Identification of Microcalcifications**" is a bonafide record of the research work done by me under the supervision of **Prof. Tessamma Thomas**, Department of Electronics, Cochin University of Science and Technology, Kerala, India. The results presented in this thesis or parts of it have not been presented for the award of any other degree.*

Kochi-22

16 August 2011

Deepa Sankar

Acknowledgements

First and foremost, I am thankful to the God Almighty for giving me the wisdom and health to complete this endeavor. During my childhood, when I see a cloud or mountain or any natural object, it was a normal sight. But now, I see these objects differently i.e. as fractals. I am indebted to the almighty for opening a world of fractals to me.

I express my heartfelt gratitude to my supervising guide Prof. Tessamma Thomas, who has supported me throughout my research work with patience and knowledge. Her constant motivation, kindness and care have helped me to cope with the difficult times during the course of my research.

I would like to offer my sincere thanks to Prof. P.R.S Pillai, Head, Department of Electronics, for providing the necessary facilities to accomplish this work.

I am extending my heartfelt gratitude to Prof. K Vasudevan, Dean, Faculty of Technology, for his immense support and encouragement.

I thank Prof. P Mohanan for his encouraging words about the research. I am thankful to other faculty members Prof. C.K Anandan, Dr. James Kurien, former faculty members Dr. K T Mathew and Dr. K.G Balakrishnan, for their kindness and co-operation. I am grateful to Dr. M.H Supriya, for her constant motivation for the timely completion of the work.

I extend my sincere gratitude to all the non-teaching staff and technical staff at the Department of Electronics especially, Mr. Ibrahimkutty, Mr.Suresh K , Mr. Francis, Ms. Vinitha Murali, Mr. Anil, Mr. Mohanan, Mr. Pradeep, Ms. Prasanna, Ms. Sudha, and former staff members Mr. Russel, Mr. Siraj, Mr. Rajeev, Mr. Nouruddin, and Ms. Bindu.

I am grateful to Dr David Peter, Principal, School of Engineering for granting me leave for the completion of the research. I appreciate the support and care given to me by former Principal, Dr Sobha Cyrus.

Dr. R Gopikakumari, Head, Division of Electronics, was always supportive during the entire course of the research and I am grateful to her, for the co-operation rendered.

Special thanks to Dr. Mini M.G, Head, Department of Electronics, Model Engineering College for her valuable help and suggestions during the course of the research. I extend my gratitude to Dr. Remadevi, Head, Department of Applied Sciences, Model Engineering College, for elaborating the mathematics behind fractals.

I am indebted to Ms. Anju Pradeep, who always cared me as an elder sister. Ms. Deepa J, Associate professor, College of Engineering, Chengannur has helped me to tackle many difficult situations during the research and I extend my sincere gratitude to her.

I remember the support given by Dr. Shahana T.K, for adjusting classes and exams during thesis writing. The advice given by my colleagues Dr. Mridula, Dr. Rekha, Dr. Binu Paul, Dr. Babita and Dr. Mythili has inspired me for the timely completion of the work.

I appreciate the help and support given by Research scholars Ullas G.K, Lindo A.O, Paulbert Thomas, Dr. Deepti Das Krishna, Cyriac M.O, Tony D, Anju P Mathews, Laila D, Sujith R, Sarin V.P, Dinesh R, Nijas M, Deepak, Nishamol M.S and Shameena V.A. Also, I am thankful to Ms. Aneesha, Research Scholar, Department of Statistics, for helping me to understand statistical analysis.

Creative discussions with my fellow researchers Praveen N, Anu Sabarish R, Ananda Resmi S, Reji A.P, Nobert Thomas Pallath, and Dr. Dineshkumar V.P have helped to shape the thesis. I recall the assistance of my students Mr. Gopinath K, Mr. Ajith Narayan, Mr. Midhun Pavithran, Mr. Gokul Jyothi and others for my research.

I extend my gratitude to the anonymous reviewers of my publications for providing valuable suggestions and motivating comments.

I am extremely grateful to my parents, who always stood behind me and provided me with all the facilities in my life. I remember the great support extended by my brother, Dr. Syam Sankar, for being there constantly, with encouragement, interest and belief in me. I am deeply indebted to my in-laws, who always understood me and took care of me whenever I needed them.

And above all, heartfelt gratitude to my Husband Sunesh, for his love, care, understanding, patience and sacrifice to achieve this goal, without whom this effort would have been worth nothing.

Deepa Sankar

Abstract

After skin cancer, breast cancer accounts for the second greatest number of cancer diagnoses in women. Currently the etiologies of breast cancer are unknown, and there is no generally accepted therapy for preventing it. Therefore, the best way to improve the prognosis for breast cancer is early detection and treatment. Computer aided detection systems (CAD) for detecting masses or micro-calcifications in mammograms have already been used and proven to be a potentially powerful tool, so the radiologists are attracted by the effectiveness of clinical application of CAD systems.

Fractal geometry is well suited for describing the complex physiological structures that defy the traditional Euclidean geometry, which is based on smooth shapes.

The major contribution of this research include the development of

- A new fractal feature to accurately classify mammograms into normal and abnormal (i) with masses (benign or malignant)
(ii) with microcalcifications (benign or malignant)
- A novel fast fractal modeling method to identify the presence of microcalcifications by fractal modeling of mammograms and then subtracting the modeled image from the original mammogram.

The performances of these methods were evaluated using different standard statistical analysis methods. The results obtained indicate that the developed methods are highly beneficial for assisting radiologists in making diagnostic decisions.

The mammograms for the study were obtained from the two online databases namely, MIAS (Mammographic Image Analysis Society) and DDSM (Digital Database for Screening Mammography).

List of Figures

1.1 Electromagnetic spectrum arranged according to energy per photon	03
2.1 Cross section of female breast	15
2.2 Two views of the breast (a) Cranio Caudal View (b) Medio Lateral View	21
2.3 Examples of Normal Breast	24
2.4 Examples of Malignant Microcalcifications (a) Original Mammograms with Malignant Microcalcifications (b) Region containing Microcalcifications (c) Some snippets of Malignant Microcalcifications	26
2.5 Examples of Benign Microcalcifications (a) Original Mammograms with Benign Microcalcifications (b) Region containing Microcalcifications (c) Some snippets of Benign Microcalcifications	26
2.6 Examples of circumscribed mass	28
2.7 Examples of Spiculated Mass	29
2.8 Benign circumscribed mass (a) Original Mammogram with Benign circumscribed mass (b) Region containing Benign circumscribed mass	29
2.9 Malignant Circumscribed mass (a) Original Mammogram with Malignant circumscribed mass (b) Region containing Malignant circumscribed mass	30
2.10 Examples of architectural asymmetry (a) Malignant (b) Benign	31
2.11 Examples of Spiculated Lesions (a) Benign (b) Malignant	31
3.1 Four pieces of the same single cauliflower is shown in 1, 2, 3 and 4.	34
3.2 Basic Construction Steps of Sierpinski triangle	37
3.3 (a) Basic steps for the construction of Cantor set, (b) Basic steps for the construction of Koch Curve, (c) Julia Set	39
3.3 (d) Mandelbrot Set	40
3.4 Schematic of Multiple reduction copy machine	40
3.5 Multiple Reduction Copy Machine with three reduction lenses and places the input figure in the form of an equilateral triangle	41
3.6 The input images to the MRCM and their corresponding output images	

obtained after three iterations	42
3.7 Cauchy Sequence: Points gets closer and closer along the sequence	44
3.8 Deformation obtained for the parallelogram	46
3.9 Image divided into non overlapping range blocks and the most suitable domain block is found by searching the entire image	50
4.1 Different classes of mammograms	56
4.2 Different Classes of Mammograms: Original and ROI taken out from the mammogram (a) Normal (b) Benign Mass (c) Malignant Mass (d) Benign Microcalcifications (e) Malignant Microcalcifications	58
4.3 Schematic for finding the FD using Differential Box counting method	66
4.4 Gray Level of Mammogram	67
4.5 Schematic for finding TPSA	72
4.6 Box Plot of the Fractal Dimension obtained using TPSA, DBCM and Blanket methods	90
4.7 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_1	93
4.8 Box Plot of the feature f_1 obtained using TPSA, DBCM and Blanket methods	96
4.9 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_2	98
4.10 Box Plot of the feature f_2 obtained using TPSA, DBCM and Blanket methods	100
4.11 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_3	102
4.12 Box Plot of the feature f_3 obtained using TPSA, DBCM and Blanket methods	105
4.13 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal	

List of Figures	iii
<hr/>	
Mammograms respectively for computing feature f_4	107
4.14 Box Plot of the feature f_4 obtained using TPSA, DBCM and Blanket methods	109
4.15 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_5	111
4.16 Box Plot of the feature f_5 obtained using TPSA, DBCM and Blanket methods	116
4.17 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_6	117
4.18 Box Plot of the feature f_6 obtained using TPSA, DBCM and Blanket methods	119
4.19. Plot of Misclassification Error for different features	121
4.20. ROC curves of Fractal dimension obtained using TPSA, DBCM and Blanket method respectively.	123
4.21 (a)-(f) Comparison of ROC curves obtained for the different fractal features	124
4.22 Flow Chart for the Classification of Mammograms using fractal feature f_6	129
5.1 Affine mapping between domain and range	143
5.2 Quad tree partitioning of the range blocks	146
5.3 Flowchart for the algorithm for fractal image modeling	148
5.4 Each Domain is divided into four blocks 1, 2, 3, and 4 each of whose mean is computed	151
5.5 Mammograms with microcalcifications (a) (c) & (e) Original Mammogram, (b) (d) & (f) Region of Interest of 64 x64	159
5.6 Examples of some starting images used during decoding	160
5.7 Sample 64x64 taken from the original image used as the starting image shown for subsequent modeling figures.	160
5.8 Modeling by Conventional Fractal Method (ROI-64x64, range-8x8)	

(a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)	163
5.9 Modeling Normal mammograms by Conventional Fractal Modeling Method	
(a) Original Image (b) Modeled Image	164
5.10 (a) PSNR (dB), (b) Mean Square Error and (c) Correlation respectively between the original and modeled image of normal mammograms for modeling using Conventional Fractal Modeling Method with ROI 64×64, range size 8×8	165
5.11 (a) Original Mammogram (b) Modeled Image(c) Detected Microcalcifications using Conventional fractal modeling method (ROI 64×64, range size 8×8)	167
5.12 Modeling by Modified Conventional Fractal Method (ROI-64×64, range-8×8)	
(a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)	169
5.13 (a) Original and (b) Modeled Normal Mammograms by modified Fractal Conventional modeling method	170
5.14 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Modified conventional fractal image coding method with ROI 64×64, range size 8×8.	171
5.15 (a) Original Mammogram (b) Modeled Image(c) Detected Microcalcifications using Modified Conventional fractal modeling method (ROI 64×64, range size 8×8).	172
5.16 Modeling by Mean Variance Method (ROI-64×64, range-8×8)	
(a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)	175
5.17 Original and Modeled Normal Mammograms by mean variance method	176
5.18 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Mean Variance fractal image coding method with ROI 64×64,	

range size 8×8 .	177
5.19 (a) Original Mammogram (b) – (c) Modeled Mammogram and Detected Microcalcifications using Mean Variance Method	178
5.20 Modeling by Entropy Method (ROI- 64×64 , range- 8×8) (a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)	181
5.21 (a) Original Normal Mammogram (b) Modeled Mammogram using Entropy Method	182
5.22 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Entropy coding method with ROI 64×64 , range size 8×8 .	183
5.23 (a) Original Mammogram (b) – (c) Modeled Mammogram and Detected Microcalcifications using Entropy Method	184
5.24 Modeling by Mass Center Fractal Method (ROI- 64×64 , range- 8×8) (a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)	186
5.25 (a) Original Mammogram (b) Modeled Mammogram by mass center method	187
5.26 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Mass Center coding method with ROI 64×64 , range size 8×8 .	188
5.27 (a) Original Mammogram (b) & (c) Modeled Mammogram and Detected Microcalcifications using mass center feature method respectively.	189
5.28 Modeling by Shade Non Shade Method (ROI- 64×64 , range- 8×8) (a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)	191
5.29 (a) Original Mammogram (b) Modeled Mammogram using Shade Non shade method	192
5.30 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Shade Non shade method with ROI 64×64 , range size 8×8 .	193

5.31 (a) Original Mammogram (b) & (c) Modeled Mammogram and Detected Microcalcifications using Shade non shade method respectively.	194
5.32 Comparison of the variation in time for the Modified conventional, Mean variance, mass center and entropy methods for different block sizes	196
5.33 Comparison of the time taken for encoding for modified fractal methods and mean variance, entropy, mass center and shade non shade methods	197
5.34 Comparison of the PSNR between the modeled image and the original image for the Modified Conventional, Mean Variance, Entropy, Mass Center and Shade-Non Shade Blocks methods of fractal Modeling	197
5.35 Comparison of the Mean Square Error between the modeled image and the original image for the Modified Conventional, Mean Variance, Entropy, Mass Center and Shade-Non Shade Blocks methods of fractal Modeling	198
5.36 Comparison of the Correlation between the modeled image and the original image for the Modified Conventional, Mean Variance, Entropy, Mass Center and Shade-Non Shade Blocks methods of fractal Modeling	198

List of Tables

3.1 Relation between Scaling factor, Number of copies and Dimension	36
4.1 No of Mammograms of each class obtained from the MIAS and DDSM Database used for the study	83
4.2 Comparison of different fractal signatures obtained for the different classes of mammograms	85
4.3 Average value of Distance D between the different classes of mammograms	85
4.4 Differential Distance D' between the different classes of mammograms	86
4.5 Comparison of the fractal dimensions obtained by TPSA, DBC and Blanket methods	87
4.6 Sample Fractal Dimension obtained for different mammograms using TPSA method	88
4.7 Mammograms correctly classified using fractal dimension computed using TPSA, DBC and Blanket methods	91
4.8 Comparison of the fractal feature f_1 obtained using TPSA, DBC and Blanket Methods	94
4.9 Sample Fractal feature f_1 values obtained for different mammograms using TPSA method	95
4.10 Mammograms Classification using Fractal feature f_1 computed using TPSA, DBC and Blanket methods	96
4.11 Comparison of the fractal feature f_2 obtained using TPSA, DBC and Blanket Methods	98
4.12 Sample Fractal feature f_2 obtained for different mammograms using TPSA method	99
4.13 Mammograms correctly classified using Fractal feature f_2 computed using TPSA, DBC and Blanket methods	101

4.14 Comparison of the fractal feature f_3 obtained using TPSA, DBC and Blanket Methods	103
4.15 Sample Fractal feature f_3 obtained for different mammograms using TPSA method	104
4.16 Mammograms correctly classified using Fractal feature f_3 computed using TPSA, DBC and Blanket methods	105
4.17 Comparison of the fractal feature f_4 obtained using TPSA, DBC and Blanket Methods	107
4.18 Sample Fractal feature f_4 obtained for different mammograms using DBC method	108
4.19 Mammograms correctly classified using Fractal feature f_4 computed using TPSA, DBC and Blanket methods	110
4.20 Comparison of the fractal feature f_5 obtained using TPSA, DBC and Blanket Methods	112
4.21 Sample Fractal feature f_5 obtained for different mammograms using TPSA method	113
4.22 Mammograms correctly classified using Fractal feature f_5 computed using TPSA, DBC and Blanket methods	114
4.23 Comparison of the fractal feature f_6 obtained using TPSA, DBC and Blanket Methods	117
4.24 Sample Fractal feature f_6 obtained for different mammograms using TPSA method	118
4.25 Mammograms correctly classified using Fractal feature f_6 computed Using TPSA, DBC and Blanket methods	120
4.46 Definition of the parameters for evaluating the detection accuracy	122
4.27 Statistical Analysis of the different fractal dimension and fractal features estimated using TPSA, DBC and Blanket methods	126
4.28 Average values of the different conventional features obtained for different classes of mammograms, its classification accuracy and Area Under ROC Curve	127

List of Tables	ix
5.1 Definition of the parameters for evaluating the detection accuracy	156
5.2 Evaluating the Modeled image obtained using conventional fractal, modeling method(ROI 64×64, range size 8×8)	166
5.3 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Conventional Fractal Coding Method	167
5.4 Evaluating the Modeled image obtained using Modified conventional fractal, modeling method (ROI 64×64, range size 8×8)	172
5.5 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Modified Conventional Fractal Coding Method (ROI 64×64, range size 8×8)	173
5.6 Time taken and the average number of domains searched for different range sizes for Modified Conventional Fractal modeling method (ROI 64×64)	174
5.7 Evaluating the Modeled image obtained using Mean Variance fractal modeling method	177
5.8 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Mean Variance Fractal Coding Method (ROI 64×64, range size 8×8)	179
5.9 Time taken and the average number of domains searched for different range sizes for Mean Variance Fractal modeling method (ROI 64×64)	179
5.10 Evaluating the Modeled image obtained using Entropy fractal modeling method	183
5.11 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Entropy Based Method (ROI 64×64, range size 8×8)	185
5.12 Variation in time taken and Average number of domains searched for different range sizes in Entropy based method (ROI 64×64)	185
5.13 Evaluation of the Modeled image obtained using Mass Center fractal modeling method with range size 8×8	189
5.14 Detection Sensitivity, Specificity, Average Number of Domains	

searched and Encoding time for Mass Center feature Method (ROI 64×64, range size 8×8)	190
5.15 Time taken and the average number of domains searched for different range sizes for Mass Center Method (ROI 64×64)	190
5.16 Evaluation of the Modeled image obtained using Shade Non shade fractal modeling method for a range size of 2× 2	194
5.17 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Shade Non Shade block Method (ROI 64×64, range size 2×2)	195
5.18 Time taken and the average number of domains searched for different range sizes in shade non shade method (ROI 64×64)	195
5.19 Comparison of different microcalcification detection methods (ROI 64×64, range 8×8, for shade non shade method, range 2×2)	199
5.20 Comparison of Microcalcification Detection %, Average Fractal encoding time, Average number of domains searched and the Average No. of Ranges coded for different microcalcification detection methods(ROI 64× 64 and range 8x8. For shade non shade: range 2×2)	200

Contents

1. Introduction

1.1 Digital Image Processing	02
1.2 History of Digital Image Processing	04
1.3 Steps in Digital Image Processing	05
1.4 Medical image processing	06
1.5 Literature survey	08
1.6 Objective of the research	09
1.7 Motivation	10
1.8 Contribution of the thesis	10
1.9 Organization of the Thesis	11

2. Epistemological study of Breast Cancer

2.1 Anatomy of female breast	14
2.2 Breast Cancer	16
2.3 Breast Imaging	17
2.3.1 Magnetic Resonance Elastography (MRE)	17
2.3.2 Electrical Impedance Spectroscopy (EIS)	18
2.3.3 Microwave Imaging Spectroscopy (MIS)	18
2.3.4 Near Infrared Spectroscopic Imaging (NIS)	18
2.3.5 Ultrasound	19
2.3.6 Magnetic Resonance Imaging (MRI)	19
2.4. Mammography	19
2.5 Finding Breast Changes	20
2.5.1 Medio Lateral Oblique and Cranio Caudal	21
2.5.2 Symptoms	22
2.5.3. Biopsy	22
2.6 Normal mammograms	23

2.7 Abnormalities in the breast	24
2.7.1 Calcifications	24
2.7.2 Masses	27
2.7.3 Architectural distortions	30
2.7.4 Spiculated lesions	31
2.8 Chapter Summary	32

3. A brief overview to the world of Fractals

3.1 From Euclidean geometry to Fractals	34
3.2 Properties of Fractals	32
3.2.1 Self Similarity	35
3.2.2 Fractal Dimension	36
3.2.3 Formation by Iteration	38
3.3 Multiple Reduction Copy Machine	40
3.4 Mathematical Foundations	42
3.4.1 The Space of Fractals	43
3.4.1.1 Metric spaces	43
3.4.1.2 Cauchy Sequence	44
3.4.2. Affine Transformations	45
3.4.3 The Contraction Mapping Theorem	46
3.4.3.1 Banach's Contraction Mapping Theorem.	47
3.4.4 Iterated Function System (IFS)	47
3.4.5. Collage Theorem	48
3.5 Fractal Image Coding	49
3.6 Literature Survey	51
3.7 Chapter Summary	54

4. Development of new Fractal features for the Classification of Mammograms into Normal, Benign and Malignant

4.1 Introduction	56
4.2 Fractal Dimension	59

4.3 Literature Survey	60
4.4 Fractal Dimension Estimation Methods	64
4.4.1. Box Counting Method	65
4.4.2 Differential Box Counting Method	65
4.4.3 Blanket Method	67
4.4.3.1 Fractal Signature	69
4.4.3.2 Differential Fractal Signatures and Distance Measurement	70
4.4.4 Triangular Prism Surface Area Method	71
4.5 Fractal Features	73
4.5.1 Fractal Feature 1 (f_1)	73
4.5.2. Fractal Features 2 and 3 (f_2 and f_3)	74
4.5.3 Fractal Feature 4 and 5 (f_4 and f_5)	74
4.5.4 Fractal Feature 6 (f_6)	75
4.6 Conventional features used for the comparison with fractal features	76
4.6.1 Statistical Descriptors	76
4.6.2 Textural features	77
4.7 Statistical Analysis	80
4.8 Implementation of Classification of Mammograms using various fractal features	82
4.8.1 Database Used	82
4.8.1.1 MIAS Database	82
4.8.2.1 DDSM Database	82
4.9 Results and Discussions	84
4.9.1 Evaluation using fractal signatures and distance measures	84
4.9.2. Evaluation of Fractal Dimension Estimated using different methods	86
4.9.2.1. Box Plot of Fractal Dimensions	88
4.9.2.2 Classification Accuracy using Fractal Dimension	90
4.9.3 Evaluation using Fractal Features $f_1 - f_6$	92
4.9.3.1 Evaluation using Fractal Feature f_1	92
4.9.3.2 Evaluation using Fractal Feature f_2	97

4.9.3.3 Evaluation using Fractal Feature f_3	101
4.9.3.4 Evaluation using Fractal Feature f_4	106
4.9.3.5 Evaluation using Fractal Feature f_5	110
4.9.3.6 Evaluation using Fractal Feature f_6	115
4.9.4 Performance Evaluation of the Features	120
4.9.4.1 Statistical analysis	122
4.9.4.2 Comparison of the performance of the fractal features with conventional features	127
4.10 Chapter Summary	130
5. Identification of Microcalcifications in Mammograms using New Fast Fractal Modeling Approaches	
5.1 Microcalcifications in mammograms	134
5.2 Fractal Modeling of Mammograms	135
5.3 Literature Survey	136
5.4 Mathematical Foundations for Fractal Image Modeling	139
5.5 Algorithm for Fractal Image Modeling	142
5.5.1. Algorithm Implementation for Mammogram modeling	145
5.5.2. Enhancing the Presence of Microcalcifications	147
5.6 Problems encountered during fractal image modeling	149
5.7 Fast Fractal Image Modeling	150
5.7.1 Mean and Variance Method	151
5.7.2 Entropy Method	152
5.7.3 Mass center Method	153
5.7.4 Shade and Nonshade Method	154
5.8 Diagnostic test accuracy for Microcalcification detection	156
5.9 Implementation of Fractal Image Modeling	157
5.9.1 Database used	157
5.9.2 Fixing of different parameters for fractal mammogram modeling	158
5.10 Results and Discussions	161
5.10.1 Implementation of Conventional Fractal modeling method	162

5.10.2 Implementation of Modified Conventional Fractal Image Modeling	168
5.10.3 Implementation of Mean Variance Method	174
5.10.4 Implementation of Entropy Method	180
5.10.5. Implementation of Mass Center Method	185
5.10.6. Implementation of Shade Non shade Method	190
5.11 Comparison of the different fractal modeling methods	195
5.12 Chapter Summary	200
6. Conclusions and Future Scope	
6.1 Thesis Highlights	204
6.2 Classification of mammograms by fractal features	204
6.3 Detection of Microcalcifications by fractal modeling	205
6.4 Suggestions for Future research	206
Bibliography	209
List of Publications	221
Index	223

Why geometry is often described as 'cold' and 'dry?'

One reason lies in its inability to describe the shape of a cloud, a mountain, a coastline, or a tree. Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line....

Nature exhibits not simply a higher degree but an altogether different level of complexity.

— Benoit Mandelbrot

Chapter 1

Introduction

Computer aided diagnosis is an important tool used by radiologists for interpreting medical images. Image processing techniques can be employed on the mammograms for the detection of breast cancer at an early stage. A brief introduction to the research is presented in this chapter. Fundamentals of digital image processing; its history and different steps involved in image processing are mentioned. The state-of-the-art technology in the mammogram image analysis is also highlighted. This chapter gives the motivation and objective behind this research. The main contributions of the present research are also highlighted.

1.1 Digital Image Processing

Sight is the most powerful of the five senses – sight, hearing, touch, smell and taste – which humans use to perceive their environment. Human beings who are blessed with eyesight start acquiring the images around them immediately after their birth. Processing, analyzing and understanding of images then become almost a routine (CeT 1998). In fact, more than 99% of the activity of the human brain is involved in processing images from the visual cortex (Dou 2009).

Image processing is the science of manipulating a picture. It covers a large number of techniques which are present in numerous applications. These techniques can enhance or distort an image, highlight certain features of an image, create a new image from portions of other images, restore an image that has been degraded during or after the image acquisition, and so on (Cra 1997).

Oxford dictionary defines image as an optical appearance produced by light from an object reflected in a mirror or refracted through a lens (Oxf 2011). Image can be formed by other types of radiant energy and devices. However, optical images are most common and are most important. The light intensity is recorded at corresponding points on a plane to form an image. The simplest kind of the intensity or the brightness image is a black and white image (Cha 2009).

An image is a two dimensional function $f(x, y)$, where x and y are spatial coordinates (Gon 2005). The amplitude of $f(x, y)$ at any pair of coordinates (x, y) is called the intensity level or gray level of the image at that point. When x , y and the amplitude f are finite discrete quantities, then it is called a digital image. Thus a digital image is an array of numbers each of which is called image elements, picture elements, pixels or pels. The field of digital image processing refers to processing digital images by means of a digital computer.

Before the advent of digital computers, machine processing of visual and other sensory images was a daunting task. During the 1970s and 1980s, the focus was on image representation using transforms and models, image filtering and restoration, still and video compression, and image reconstruction. Although mainframes were

originally used, affordable minicomputers became popular. This progress in computer hardware as well as in image acquisition and display devices enabled image-processing research groups to emerge around the world. Since the mid-50s, powerful workstations and personal computers have made desktop or even laptop image-processing research and technology possible. Later, with advances in computing, memory, and image-sensing technology, techniques developed for image enhancement, still and moving image compression, and image understanding gave this field a solid base of practical applications (CeT 1998).

More recently, technology has tremendously extended the possibilities for visual observation. Photography makes it possible to record images objectively, preserving scenes for later, repeated, and perhaps more careful, examination. Telescopes and microscopes greatly extend the human visual range, permitting the visualization of objects of vastly differing scales. Technology can even compensate for inherent limitations of the human eye. The human eye is receptive to only a very narrow range of frequencies within the electromagnetic spectrum (Fig. 1.1)

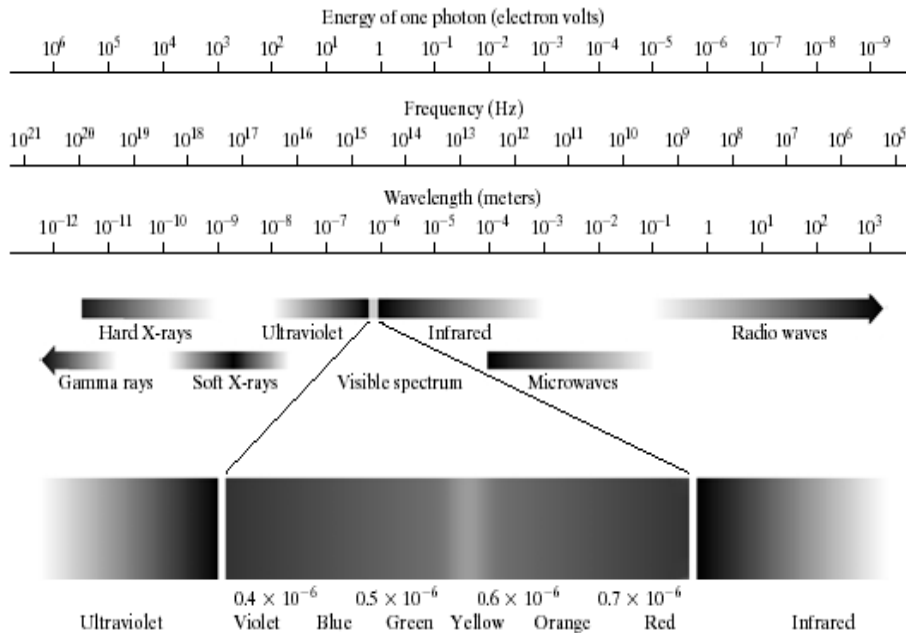


Fig 1.1 Electromagnetic spectrum arranged according to energy per photon

Nowadays, there are sensors capable of detecting electromagnetic radiation outside this narrow range of “visible” frequencies, ranging from γ -rays and x-rays, through ultraviolet and infrared, to radio waves. Today, there is almost no area of technical endeavor that is not impacted in some way or other by digital image processing (Dou 2009).

1.2 History of Digital Image Processing

The history of digital images is quite young. First of the digital images appeared in the earlier 1920s (Gon 2005). The first application was in the news paper industry, when pictures were sent by submarine cable between London and New York. The introduction of Bartlane cable picture transmission system, in the early 1920s, helped to reduce the time taken to transmit across the Atlantic from more than a week to less than 3 hours. These pictures initially had only 5 distinct gray levels, but increased to 15 gray levels by the 1929.

As digital computers were not involved for the creation, these examples cannot be considered as part of digital image processing. The first computers to carry out meaningful image processing tasks appeared in the early 1960s. Since then there was no looking back (Gon 2005).

In 1964, pictures of the moon was transmitted by *Ranger 7*, which were the first images taken by the U.S spacecraft (Cra 1997). Over the years, NASA had plenty of images to process. The Ranger spacecraft provided hundreds of images of the lunar surface. The Surveyor 7 spacecraft returned 21,038 television images of its landing site on the moon. The Mariner 4, launched in 1964, returned 22 digital images of Mars. The Viking missions started in 1975 and they provided over 100,000 images of Mars. The Voyager mission, in 1977, launched two spacecraft that returned a wide range of imagery of the outer planets: Saturn, Uranus, Neptune, and Jupiter.

In late 1960s and early 1970s in parallel with space applications, image processing techniques were used in medical imaging, remote earth resources and astronomy. From the 1960s until the present, the field of image processing has grown

vigorously. They have a broad range of applications in interpreting images in industry, medicine biological sciences, and physics. The typical problems in machine perception includes automatic character recognition, industrial machine vision for product assembly and inspection, military recognizance, automatic processing of finger prints, screening of X-rays and blood samples and machine processing of aerial and satellite imagery for weather prediction and environmental assessment (Gon 2005).

1.3 Steps in Digital Image Processing

In all the applications of image processing, *image acquisition* is the first step. Numerous electromagnetic and some ultrasonic sensing devices are frequently arranged in the form of a 2-D array. The response of each sensor is proportional to the light energy falling onto the surface of the sensor. Generally image acquisition stage involves preprocessing like scaling (Gon 2005).

The simplest and most appealing areas of digital image processing are the *image enhancement*. This is a subjective approach. The goal is to process the image so that the result is more suitable than the original image for a specific application. The word specific is important because the methods for enhancing one kind of images may not be suitable for another kind, eg. X-ray images and space craft images.

Image restoration attempts to reconstruct or recover an image that has been degraded by using an a priori knowledge of degradation phenomenon and is based on mathematical and probabilistic models of image degradation. This includes deblurring of images degraded by the limitations of a sensor or its environment, noise filtering, and correction of geometric distortion or nonlinearities due to sensors (Jai 1989).

Image analysis techniques require extraction of certain features that aid in the identification of the object. *Segmentation* techniques are used to isolate the desired object from the scene so that measurements can be made on it subsequently. *Segmentation* partitions the image into its constituent parts or objects. The level to which the subdivision is carried depends on the problem being solved. *Representation*

and description almost follow the output of a segmentation stage, which is usually raw pixel data, constituting the boundary of the region, i.e. a set of pixels separating one region from another or all the points in it. In either case, converting data to a suitable form for computer processing is necessary. *Description* is also called feature selection. It deals with extracting the attributes that result in some quantitative information of interest or is basic for differentiating one class of object from another.

Recognition is a process that assigns a label to an object, based on its descriptors.

1.4 Medical image processing

The advent of medical imaging is one of the milestones in the progress of medical science. Medical imaging systems detect different physical signals arising from a patient and produce images. It serves as a beneficial tool for the medical practitioners during diagnosis of ailments.

An imaging modality is an imaging system which uses a particular technique for producing the image. Some of these modalities use ionizing radiation, radiation with sufficient energy to ionize atoms and molecules within the body and others use non ionizing radiation. Ionizing radiation in medical imaging comprises x-rays and γ -rays, both of which need to be used prudently to avoid serious damage to the body and to its genetic material. Non-ionizing radiation like, ultrasound and radio frequency waves, on the other hand, does not have the potential to damage the body directly and the risks associated with its use are considered to be very low.

The application of image processing techniques to medical imaging has made the results accurate and reliable. In many cases it is possible to eliminate the necessity for invasive surgery, thus avoiding trauma to the patient as well as inevitable element of risk. One of the early applications of image processing in the medical field is the enhancement of conventional radiograms. When converted to digital form, it is possible to remove noise element from x-ray images thereby enhancing their contrast. This aids interpretation and removes blurring caused by unwanted movement of the

patient. This form of representation also enables the physicians to measure the extent of tumors and other significant features accurately.

In medical imaging, the perfect diagnosis and/or assessment of a disease depends on both image acquisition and image interpretation. The advances in medical quality compliance regulations, image detector systems and computer technology have tremendously increased the role and contribution of radiology to medical diagnosis. For example, a major contributor to the improvement in medical imaging has been cross-sectional imaging (e.g., X-ray computed tomography (CT) and Magnetic Resonance Imaging (MRI)), which depends greatly on computer power and data storage capabilities, and produces many three-dimensional (3-D), high-quality images for interpretation.

The image interpretation process, however, has only recently begun to benefit from computer technology. Most interpretations of medical images are performed by radiologists; however, image interpretation by humans is limited due to the nonsystematic search patterns of humans, the presence of structure and noise (camouflaging normal anatomical background) in the image, and the presentation of complex disease states requiring the integration of vast amounts of image data and clinical information.

Computer Aided Diagnosis (CAD), defined as a diagnosis made by a radiologist who uses the output from a computerized analysis of medical images as a “second opinion” in detecting lesions, assessing extent of disease, and making diagnostic decisions, is expected to improve the interpretation component of medical imaging. With CAD, the final diagnosis is made by the radiologist. Computerized image analysis has been applied mainly to medical imaging techniques such as X-ray, sonography, and Magnetic Resonance Imaging (Gig 2001).

X-ray imaging is a transmission-based technique in which X-rays from a source pass through the patient and are detected either by film or an ionization chamber on the opposite side of the body.

Breast cancer is one of the common cancer forms affecting women worldwide. Each year, more than 180,000 new cases of invasive breast cancer are diagnosed and more than 40,000 women die from the disease (Nas 2001). Early detection is the only

hope for reducing the burden of disease due to breast cancer. Clinical data show that women diagnosed with early-stage breast cancers are less likely to die of the disease than those diagnosed with more advanced stages of breast cancer.

X-ray mammography has been able to detect cancer at an earlier stage, reducing disease specific mortality. Mammograms are particularly difficult to interpret for women with dense breast tissue, as dense tissues interfere with the identification of abnormalities associated with tumors. Screening mammograms produces a large number of mammograms which are generally normal ones. Thus there is a chance that radiologists, who have a huge case load, make mistake while taking decision.

The major categories of error are due to poor radiographic technique, absence of radiographic criteria of cancer, obvious oversight by the radiologist and lack of recognition of subtle radiographic sign (Mar 1979). To cater to this problem, different image processing techniques are applied for the Computer Aided Diagnosis in digital mammogram, which help the radiologists in taking decisions.

1.5 Literature survey

In mammography, the contrast between the soft tissues of the breast is intrinsically small making the interpretation of a mammogram difficult. Also, a relatively small change in the mammographic structure can indicate the presence of a malignant breast tumor.

Polokowski *et.al.* (Pol 1997) developed a new model-based vision (MBV) algorithm to find out regions of interest (ROI's) corresponding to masses in digitized mammograms and to classify the masses as malignant/benign.

Sameti *et. al* (Sam 2009) introduced a stepwise discriminant analysis with six features to distinguish between the normal and abnormal regions. The best linear classification function resulted in a 72% average classification.

A dense to sparse microcalcification clusters grouping method based on distance independent of size, shape and orientation of real clusters was proposed by Mao *et. al.* (Mao 1998).

J. Tang *et.al* (Tan 2009) proposed a new image enhancement technology based on a multiscale contrast measure in the wavelet domain for radiologists, for screening the mammograms. Peng *et.al* (Pen 2009) employed a Stochastic Resonance (SR) noise based detection algorithm to enhance the detection of microcalcifications in mammograms.

The left–right (bilateral) asymmetry in mammograms was analyzed based on the detection of linear directional components by using a multiresolution representation based on Gabor wavelets. This gave an average classification accuracy of 74.4% (Fer 2001).

Faye *et.al* (Fay 2009) decomposed mammogram images using Daubechies 3 wavelet function and the corresponding coefficients extracted were used to differentiate between normal and abnormal mammograms and to classify the abnormal ones into benign or malignant tumors with an average classification accuracy of 98%.

The gradient-based features and texture measures based on gray-level co-occurrence matrices (GCMs) were used for the classification of mammographic masses as benign or malignant by Mudigonda *et. al.* (Mud 2000). Their method produced a benign versus malignant classification of 82.1%, with an area Az of 0.85 under the receiver operating characteristics (ROC) curve.

Context features that represent suspiciousness of normal tissue were developed for the detection of malignant masses in mammograms (Hup 2009). The Free response receiver operating characteristic (FROC) curves were computed for feature sets including context features and a feature set without context. Results show that the mean sensitivity in the interval of 0.05–0.5 false positives/image increased more than 6% when context features were added.

Detailed literature reviews on appropriate fields are presented from chapter 3 onwards.

1.6 Objective of the research

The objective of this research work is

- To classify mammograms into normal and abnormal. Abnormalities include masses and microcalcifications which are benign and malignant.
- After classification, mammograms with microcalcifications are considered. The region containing microcalcifications in the mammograms are identified.

Fractal based methods are employed in the present work.

1.7 Motivation

Breast cancer is one of the leading causes of mortality among women. At present, India reports around 100,000 cases of breast cancer annually. According to a study by International Agency for Research on Cancer (IARC), a branch of World Health Organization (WHO), there will be approximately 250,000 new cases of breast cancer in India by 2015 (Bre 2011). In the United States, one in eight women is affected by breast cancer, which kills more women than any cancer except lung cancer (ACS 2008). But early detection of breast cancer can help in reducing the mortality rate by 30%.

The breast parenchymal and ductal patterns are highly self similar, which is the basic property of fractals. Therefore; fractal analysis can be applied in mammograms.

1.8 Contribution of the thesis

The main contributions of this research include:

- The development of a new fractal feature which gave high classification accuracy for the efficient classification of mammogram into normal and abnormal and its subclasses.
- The development of a new fast fractal based mammogram modeling method with improved detection score, for the identification of microcalcifications, which are early indication of breast cancer.

1.9 Organization of the Thesis

Chapter 2 deals with the description and the imaging modalities for the detection of breast cancer. The different classes of mammograms like masses, microcalcifications etc are also presented in this chapter.

Chapter 3 is dedicated to the description of fractals. The fundamental properties and mathematical background are detailed here.

Chapter 4 presents the classification method based on fractal features. The basic property of fractal dimension was used for the classification. The three different fractal dimension estimation methods like the differential box counting method, blanket method and triangular prism surface area method are discussed. The six fractal features derived from these methods and the distance measures used to differentiate between the different classes of mammograms are also presented in this chapter.

Chapter 5 deals with the extension of fractal image modeling for the detection of microcalcifications. Here the self similarity property of fractals is exploited. The time taken for the fractal image coding is too large and four methods based on mean and variance, entropy, mass center and shade – non shade blocks were introduced. This considerably reduced the encoding time as well as increased the microcalcification detection accuracy.

The summary and conclusions based on the present work are given in *Chapter 6*. A brief description on the future prospects and possibility of the continuation of the present work are also included in this chapter.

Chapter 2

Epistemological study of Breast Cancer

Oncologists world over are concerned about the high growth rate of breast cancer cases among womanhood. As this research is intended to develop techniques to detect breast cancer at an early stage, a medical perspective of breast cancer is presented in this chapter. Anatomy of female breast is explained in the beginning. The current most popular and cost effective breast imaging modality is the x ray images of breast called mammograms. The symptoms of breast cancer and biopsies required are explained. Different abnormalities that are visible in mammograms are also mentioned.

This research work is aimed at developing a new computer aided method for the early detection of breast cancer. Therefore this chapter provides a brief insight into the medical perspective of breast cancer.

Breast cancer is one of the best-studied human tumors, but it remains poorly understood. Although it is certain that, breast cancer is the result of DNA alterations (damage or mutation) that lead to uncontrolled cell proliferation, the actual etiology of breast cancer remains obscure. A basic understanding of the anatomy and histology of the breast is important for an understanding of the pathologic processes that occur and are helpful for image interpretation (Kop 2007).

2.1 Anatomy of female breast

The breast generally refers to the front of the chest and medically specific to the mammary gland (Med 2009). The breast is a mass of glandular, fatty, and fibrous tissues positioned over the pectoral muscles of the chest wall and attached to the chest wall by fibrous strands called Cooper's ligaments. A layer of fatty tissue surrounds the breast glands and extends throughout the breast. The fatty tissue gives the breast a soft consistency (Bel 2009). The cross sectional view of female breast is given in fig 2.1.

Each breast has 15 to 20 sections, called lobes that are arranged like the petals of a daisy. Each lobe has many smaller lobules, which end in dozens of tiny bulbs that can produce milk. The lobes, lobules, and bulbs are all linked by thin tubes called ducts. Toward the nipple, each duct widens to form a sac (ampulla). During lactation, the bulbs on the ends of the lobules produce milk. Once milk is produced, it is transferred through the ducts to the nipple (Bel 2009). These ducts lead to the nipple in the center of a dark area of the skin called the areola (OSU 2011).

There are no muscles in the breast, but muscles lie under each breast and cover the ribs. Each breast also contains blood vessels and vessels that carry lymph. The lymph vessels lead to small bean-shaped organs called lymph nodes, clusters of which are found under the arm, above the collarbone, and in the chest, as well as in many other parts of the body (OSU 2011).

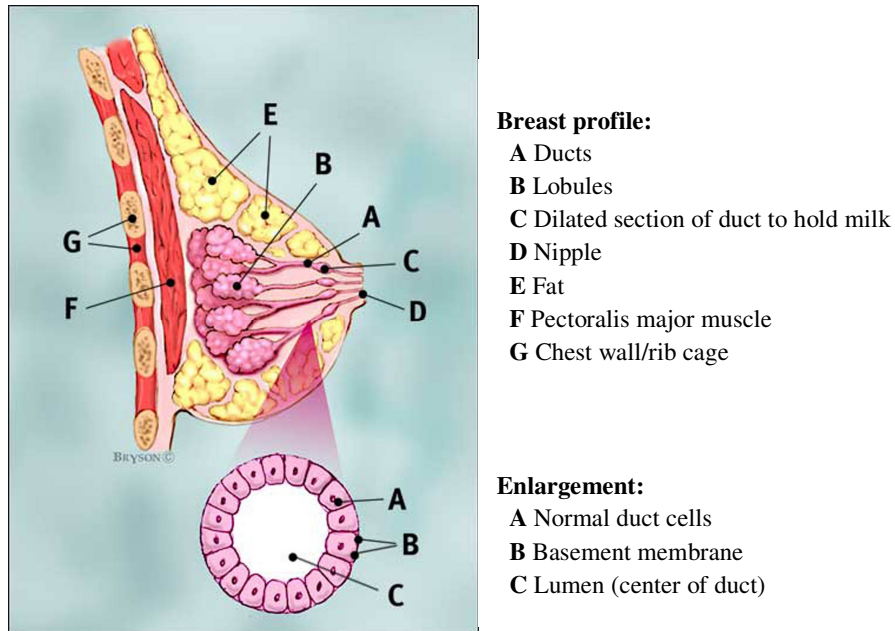


Fig.2.1 Cross section of female breast

(Courtesy: http://www.breastcancer.org/pictures/breast_anatomy/)

The shape and appearance of the breast undergo a number of changes as a woman ages. In young women, the breast skin stretches and expands as the breasts grow, creating a rounded appearance. Young women tend to have denser breasts (more glandular tissue) than older women (Bel 2009).

A woman's breasts are rarely balanced (symmetrical). Usually, one breast is slightly larger or smaller, higher or lower, or shaped differently than the other. The size and characteristics of the nipple also varies from one woman to another. During each menstrual cycle, breast tissue tends to swell from changes in the body's levels of estrogen and progesterone. The milk glands and ducts enlarge, and in turn, the breasts retain water (Bel 2009).

During pregnancy, a variety of breast changes occur. Typically, breasts become tender and the nipples become sore after a few weeks of conception. The

breasts also increase in size very quickly. The nipples also become larger and more erect as they prepare for milk production.

The breasts' glandular tissue, which has been kept firm so that the glands could produce milk, shrinks after menopause and is replaced with fatty tissue. The breasts also tend to increase in size and sag because the fibrous (connective) tissue loses its strength. It is easier for the radiologists to detect cancer on older women's mammogram films, since the breast becomes less dense after menopause. The abnormalities will be more visible as breast is less dense. Since the risk of breast cancer increases with age, all women should undergo annual screening of mammograms after the age of 40, and continue monthly breast self-exams and physician-performed clinical breast exams every year.

2.2 Breast Cancer

Cancer begins in the cells which are the basic building blocks of the body and it is named after the place from where it originates. Normally, body forms new cells as needed, replacing old cells that die (NLM 2010). This is a normal, controlled process. However, there is a chance that this orderly process could be disturbed and cells begin to reproduce in an abnormal way (Min 2003). New cells grow even when it is not needed. These extra cells can form a mass called a tumor (NLM 2010).

Tumors can be benign or malignant. Benign tumors remain similar to the tissue of their origin. Generally, benign tumors are not cancerous while malignant ones are. Cells from malignant tumors can invade nearby tissues. They can also break away and spread to other parts of the body (Kop 2007).

When a tumor spreads to other parts of the body and grows, invading and destroying other healthy tissues, it is said to have metastasized. This process is called metastasis, and the result is a serious condition that is very difficult to treat (MNT 2004).

Alterations of considerable extent are present in the mammary duct epithelium of each breast which contains a primary carcinoma, whether infiltrating or non infiltrating (Gal 1969).

At present, high quality mammography is the diagnostic method with the proven highest accuracy in finding early breast cancer at the lowest cost–benefit and harm–benefit ratios (Tab 2003). If it is detected at an early stage, the survival rate of the patients can be increased.

On mammogram films, breast masses, including both non-cancerous and cancerous lesions, appear as white regions. Fat appears as black regions on the films. All other components of the breast (glands, connective tissue, tumors, calcium deposits, etc.) appear as shades of white on a mammogram. In general, for younger woman the breasts are denser. As woman ages, her breasts become less dense and the space are filled with fatty tissue shown as dark areas on mammography x-rays (Bel 2009).

If two or more readers review these mammogram images, it reduces the failure to perceive an abnormality. Unfortunately, two radiologists reviewing every image are not practical. Nevertheless, it is a way to reduce the chance of overlooking a cancer on a mammogram. Computer-aided diagnosis (CAD) comes as a help in this problem. Double reading, may mean a review by two readers to reduce errors of perception, or it may be considered as double interpretation, where the second reader may decide on the concerns raised by the first reader as warranted or not (Kop 2007).

The most popular methods for interpreting mammograms presented in the Atlas of mammography by Tabar (Tab 2001) are discussed in the next section.

2.3 Breast Imaging

Different breast imaging (Pau 2005) modalities which help in the diagnosis of breast cancer are discussed in this section.

2.3.1 Magnetic Resonance Elastography (MRE)

In this technique, mechanical vibrations are applied to the breast's surface that propagates through the breast as a three-dimensional, time-harmonic spatial displacement field varying locally with the mechanical properties of each tissue region. These data are used to optimize a Finite Element (FE) model of the breast's

three-dimensional mechanical property distribution by iteratively refining an initial estimate of that distribution until the model predicts the observed displacements as closely as possible.

2.3.2 Electrical Impedance Spectroscopy (EIS)

EIS passes small AC currents through the pendant breast by means of a ring of electrodes placed in contact with the skin. Magnitude and phase measurements of both voltage and current are made simultaneously at all electrodes. The observed patterns of voltage and current are a function of both the signals applied and of the interior structure of the breast. EIS is referred to as electrical impedance spectroscopy because AC currents can be applied to the breast at a wide range of frequencies.

2.3.3 Microwave Imaging Spectroscopy (MIS)

Like EIS, MIS interrogates the breast using EM fields. It differs in using much higher frequencies (300–3000 MHz). In this range it is appropriate to treat EM phenomena in the breast in terms of wave propagation rather than voltages and currents. The technologies and mathematics used in EIS and MIS are, therefore, divergent, despite the fact that both exploit EM interactions in tissue.

2.3.4 Near Infrared Spectroscopic Imaging (NIS)

In NIS, a circular array of optodes (in this case, optical fibers transcribing infrared laser light) is placed in contact with the pendant breast. Each optode in turn is used to illuminate the interior of the breast, serving as a detector when nonactive. A two or three-dimensional FE model of the breast's optical properties is iteratively optimized until simulated observations based on the model converge with observation.

2.3.5 Ultrasound

An ultrasound device that uses high frequency sound waves which bounce off tissues and echoes are converted to pictures. The pictures produced show whether a lump is solid or filled with fluid. This exam may be used along with a mammogram (Bel 2009).

2.3.6 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) uses a powerful magnet linked to a computer (Bel 2009). MRI makes detailed pictures of breast tissue. MRI may also be used along with a mammogram.

The most common breast imaging modality is the mammogram and is explained in the next section.

2.4. Mammography

Mammography is a radiographic examination that is specially designed for detecting breast pathology. It is the single most important technique in the investigation of breast cancer. It can detect disease at an early stage when therapy or surgery is most effective (Dou 2009), (Bic 2010). A mammogram is a picture of the breast that is made by using low-dose x-rays (Bel 2009).

X-ray mammography is one of the most challenging areas in medical imaging. It is used to distinguish subtle differences in tissue type and detect very small objects, while minimizing the absorbed x-ray dose to the breast. Since the various tissues comprising the breast are radiologically similar, the dynamic range of mammograms is low. Most modern x-ray units use molybdenum targets, instead of the usual tungsten targets, to obtain an x-ray output with the majority of photons in the 15–20 keV range (Dou 2009).

To see lesions in dense fibro glandular tissue, the x-ray beam should be sufficiently energetic to penetrate these tissues. The American College of Radiology (ACR) recommends that the least penetrated tissues on a film/screen mammogram (the whitest areas) measure 1.0 or higher on a densitometer. This is a level at which structures can be seen through the dense (white) portions of the mammograms. There must also be sufficient penetration on a digital mammogram so that structures are also visible in the least penetrated areas (Kop 2007), (Bus 2002).

However the interpretation of screening mammograms is a repetitive task involving subtle signs, and suffers from a high rate of false negatives (10–30% of women with breast cancer are falsely told that they are free of the disease on the basis of their mammograms (Mar 1979), and false positives (only 10–20% of masses referred for surgical biopsy are actually malignant).

2.5 Finding Breast Changes

Screening is looking for cancer before a person has any symptoms (Bel 2009). This can help to find cancer at an early stage. When cancer is found early, it is easier to treat. By the time symptoms appear, cancer may have begun to spread. Three tests are commonly used to screen for breast cancer:

- Mammogram: Taking the x-ray of the breast.
- Clinical breast exam (CBE): A clinical breast exam is an examination of the breast by a doctor or other a health professional. The doctor will carefully feel the breasts and under the arms for lumps or anything else that seems unusual.
- Breast self-exam (BSE): Breast self-exam refers to examination to check their own breasts for lumps or anything else that seems unusual.

While screening mammography attempts to identify breast cancer in the asymptomatic population, diagnostic mammography are performed to further evaluate abnormalities such as a palpable mass in a breast or suspicious findings identified by screening mammography. In screening mammography, two x-ray images of each breast, in the mediolateral oblique and craniocaudal views are

routinely acquired. Early detection of cancer gives the patient the best chance to have the disease successfully treated. The American Medical Association, American Cancer Society, and American College of Radiology have recommended yearly mammogram for women, after the age of forty.

2.5.1 Medio Lateral Oblique and Cranio Caudal

Screening mammography involves taking two views of the breast, from above (cranial-caudal view, CC) and from an oblique or angled view (mediolateral-oblique, MLO) (NLM 2010) and is shown in fig. 2.2.

The mediolateral oblique view (MLO) is taken from an oblique or angled view. During routine screening mammography, the MLO view is preferred over a lateral 90-degree projection because more of the breast tissue can be imaged in the upper outer quadrant of the breast and the axilla (armpit).

The cranio-caudal view (CC) images the breast from above. With the CC view, the entire breast parenchyma (glandular tissue) should be depicted. The fatty tissue closest to the breast muscle should appear as a dark strip on the x-ray and behind that it should be possible to make out the pectoral (chest) muscle. The nipple should be depicted in profile.

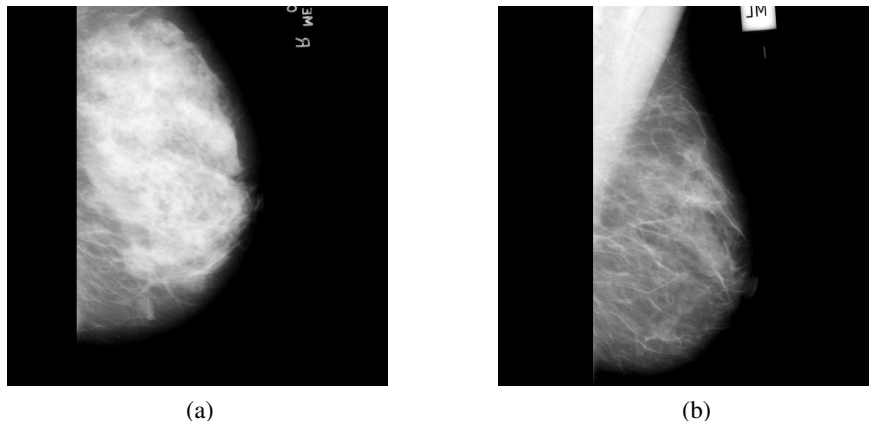


Fig 2.2 Two views of the breast (a) Cranio Caudal View (b) Medio Lateral View

2.5.2 Symptoms

Symptoms of breast cancer can vary depending on the stage it is in. There may not be visible symptoms in the early stages of breast cancer. Usually there is no pain or any other sign associated with breast cancer.

As the cancer grows it can cause changes in the breast (Bel 2009). Some of them are:

- A lump in or near the breast or under arm
- Thick or firm tissue in or near the breast or under arm
- Nipple discharge or tenderness
- A nipple pulled back (inverted) into the breast
- Itching or skin changes such as redness, scales, dimples, or puckers
- A change in breast size or shape

Since risk of breast cancer increases with age, monthly breast self-exams and physician-performed clinical breast exams every year are preferred (Bel 2009).

Once a change in the breast is sensed, biopsy and histological examination are generally necessary to determine whether the abnormality identified by the imaging methods is benign and harmless or malignant and life threatening.

2.5.3. Biopsy

If a breast lump is suspicious, a sample of tissue or fluid must be taken and tested by a pathologist (Bel 2009). There are four methods for biopsy:

Fine needle aspiration (FNA)

A very fine needle is used to remove out some fluid and /or cells, which will be examined under a microscope. FNA is the fastest and easiest method of breast biopsy, and the results are rapidly available.

Core needle biopsy

The needle used during core needle biopsy is larger than the needle used with FNA. The surgeon puts this needle through the skin and into the suspicious regions in the mammograms which are palpable and non palpable tissue samples are taken. The sample is then sent for lab tests.

Vacuum-assisted biopsy (Mammotome or minimally invasive breast biopsy)

This relies on the suspicious regions identified by the stereotactic mammography or ultrasound imaging. The computer coordinates will help the physician to guide the needle to the correct area in the breast. Vacuum-assisted biopsy is a minimally invasive procedure that allows the removal of multiple tissue samples. After the biopsy is complete, the tissue samples will be sent to the pathology laboratory for diagnosis.

Open surgical (excisional or incisional)

During an excisional surgical biopsy, the surgeon will attempt to completely remove the area of concern (lesion), often along with a surrounding margin of normal breast tissue. An incisional surgical biopsy is similar to an excisional biopsy except that the surgeon removes only the part of the breast lesion and is usually only performed on large lesions.

2.6 Normal mammograms

Unfortunately, there is no normal appearance on a mammogram that can be memorized. What constitutes “normal” varies within a wide spectrum. In addition, the appearance of the breast differs during pregnancy and in the post partum period. This spectrum is due to the difference in the breast composition. Snippets of normal mammograms are shown in fig 2.3.

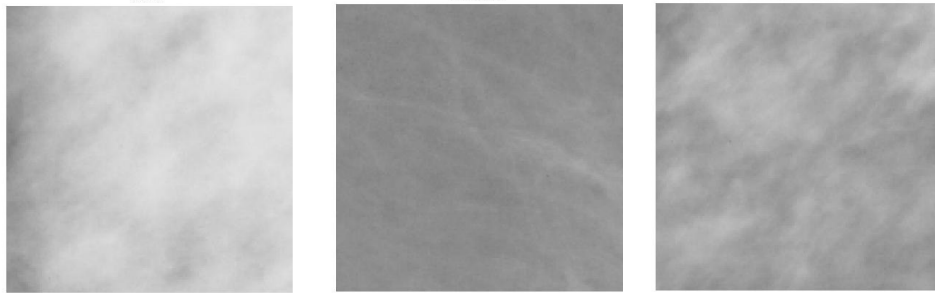


Fig. 2.3 Examples of Normal Breast

A breast with a high composition of adipose tissue will appear darker on a mammogram than a breast with a high composition of connective tissue stroma (tighter). Other normal variations of breast tissue include asymmetric patterns and asymmetric size. Although the breasts usually develop symmetrically, differences in the symmetry of the breast tissue patterns or breast size are not necessarily abnormal. Without other indices of an abnormal process, such asymmetry may simply be a development phenomenon.

2.7 Abnormalities in the breast

This section describes the different abnormalities that are seen in the mammograms.

2.7.1 Calcifications

A calcification is a deposit of the mineral calcium in the breast tissue. Calcifications appear as small white spots on a mammogram (Bel 2009). There are two types:

Macrocalcifications are large calcium deposits often caused by age. These are not usually cancerous.

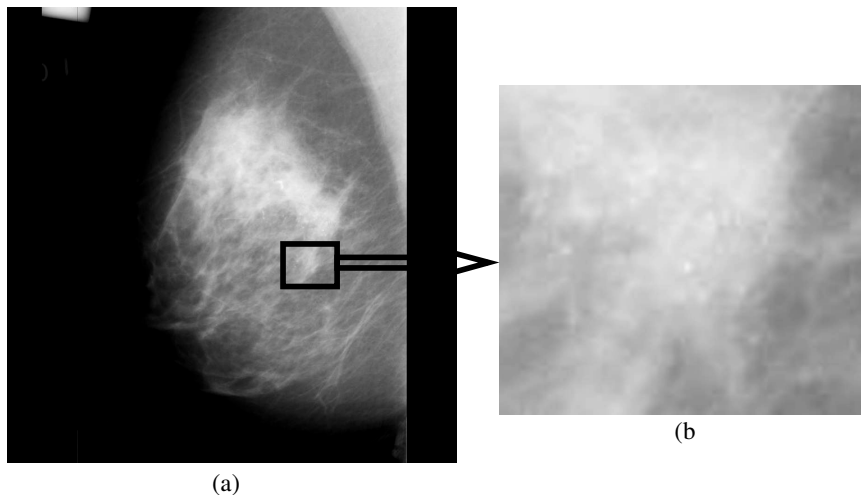
Microcalcifications are tiny specks of calcium that may be found in the area of rapidly dividing cells. If they are found grouped together in a certain way, it may be a sign of cancer.

Calcium in the diet does not create calcium deposits (calcifications) in the breast. A cluster is typically defined as 3 to 5 microcalcifications within 1 square centimeter region. Up to 50% of malignant masses demonstrate clustered microcalcifications and in a number of cases the clusters are the only signs of malignancy.

Suspicious calcifications occur in about one third of breast cancers and may develop prior to the invasive phase of tumor growth (in situ cancer) when cancer is most curable. Calcium deposits are easy to be seen in x-rays because they are much denser (have higher x-ray stopping power) than all types of soft tissues in the breast. Calcifications associated with cancer are usually very small. However, calcifications commonly occur in benign breast processes where they may be confused with cancer. Examples of malignant and benign microcalcifications are shown in fig. 2.4 and 2.5 respectively.

Malignant microcalcifications vary extremely in form, size, density and number. They are usually clustered within one area of the breast, often within one lobe. They are two types: granular and casting.

Pleomorphic or heterogeneous calcifications (granular): These are tiny calcifications with dot like or elongated shape and are innumerable. They are varying in size, usually less than 0.5mm. Fine and/or branching (casting) calcifications: these are thin irregular calcifications that appear linear, but are discontinuous and under 0.5mm in width. They are often associated with cancer and clearly merit immediate biopsy.



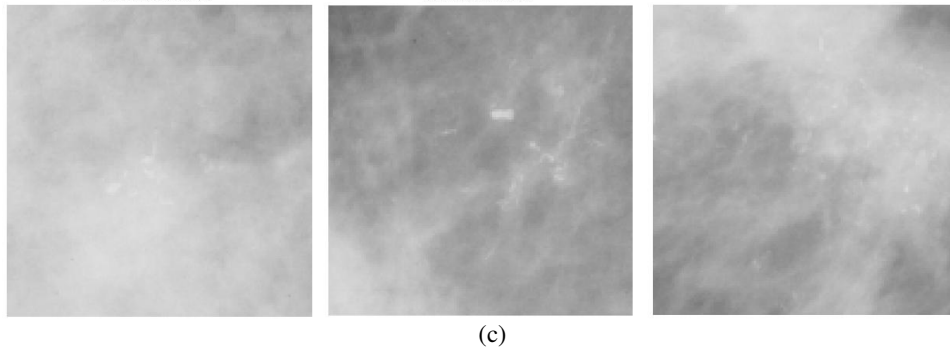


Fig 2.4 Examples of Malignant Microcalcifications (a) Original Mammograms with Malignant Microcalcifications (b) Region containing Microcalcifications (c) Some snippets of Malignant Microcalcifications

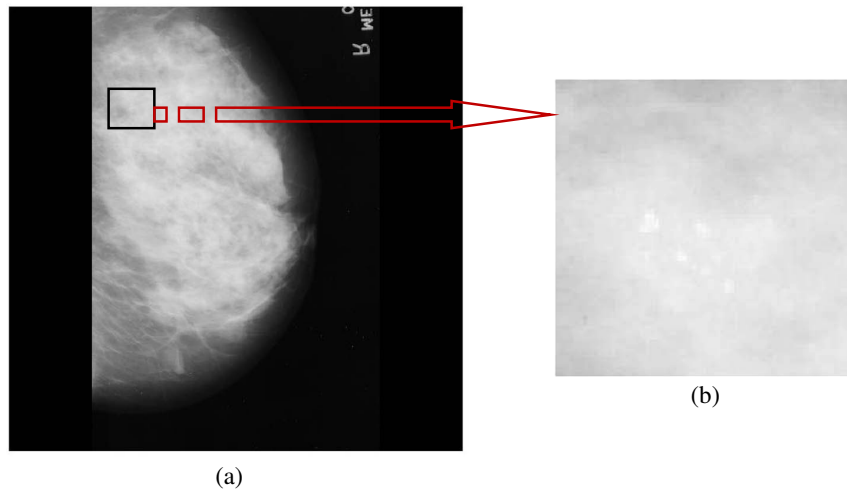


Fig. 2.5 Examples of Benign Microcalcifications (a) Original Mammograms with Benign Microcalcifications (b) Region containing Microcalcifications (c) Some snippets of Benign Microcalcifications

Benign microcalcifications are characterized by homogeneous shape, density, sharp outline or radio lucent density. They are usually larger than that associated with malignancy. They are coarser, often round with smooth margins and they are easily seen.

Skin calcifications: these are typically dense smooth and lucent centers (less dense in center than margin) that are pathognomonic (appearance is always benign). They are situated in the skin, resulting from calcium deposits in hair follicles and are more common in the center of the chest at the inner edge of the breast.

Vascular calcifications: they are parallel paired tracks or linear tubular calcifications that are clearly associated with small arteries.

Coarse or pop corn like calcifications: Rounded groups of coarse calcifications develop in an involution fibro adenoma. When completely developed the appearance is reliable, but during early phases of development, calcifications in fibro adenomas may be suspicious.

Large rod shaped calcification: these are benign calcifications forming continuous rods that may occasionally branch. They are usually more than 1mm in diameter and may have lucent center, if calcium surrounds rather than fills an enlarged duct.

Round calcifications: they are smooth dense and round calcifications with size less than 1mm.

Spherical or lucent centered calcifications: These are benign calcifications that range from under 1mm to over a centimeter. These deposits have smooth surfaces, are round or oval and tend to have a lucent center. They arise from areas of fat necrosis, calcified ducts debris and occasional *fibro adenoma* of a duct involved irregularly by breast cancer.

2.7.2 Masses

The presence of a localized collection of tissue represents a mass. By ACR BIRADS definition, a mass is a space occupying lesion seen in 2 different projections (X-rays point of view). When an apparent collection is seen in only one view, it is

referred to as a mammographic density. Although the density may be of a mass perhaps obscured by overlying glandular tissue on other views, it may be nothing more than several overlapping normal areas. When a density is seen in only one view, additional views must be done to confirm the presence of a mass.

Circumscribed masses have a distinct border and are typically circular in shape. High density radio lucent /radio opaque combined masses are almost always benign.

Halo and capsules are characteristics of benign masses with rare exceptions. A halo is a narrow radiolucent ring or a segment of a ring around periphery of a tumor. A capsule is a thin curved radio opaque line that is seen only when it surrounds tumors containing fat. A cyst with smooth borders and orienting in the direction of the nipple following the trabecular structure of the breast also indicates a benign lesion. Contour density, shape, orientation and size of the mass are important factors to be considered when analyzing a visible mass. General shape of a mass is relatively non specific since both benign and malignant processes tend to arise from one spot and grow circumferentially. Examples of circumscribed and spiculated masses are shown in fig. 2.6 and 2.7 respectively.

Round and oval shapes are associated with benign processes in part because they imply a well circumscribed region, a benign sign considered in the margins section.

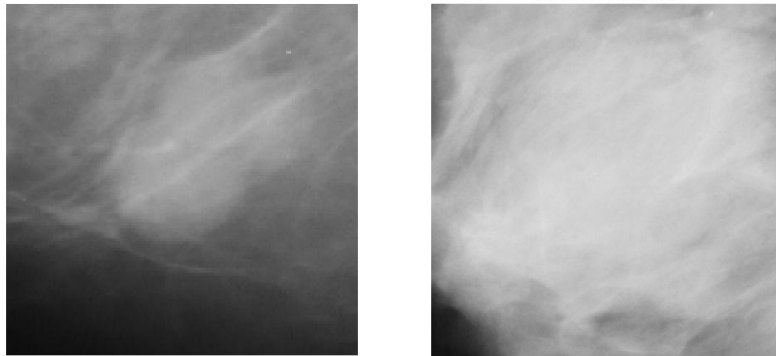


Fig 2.6 Examples of circumscribed mass

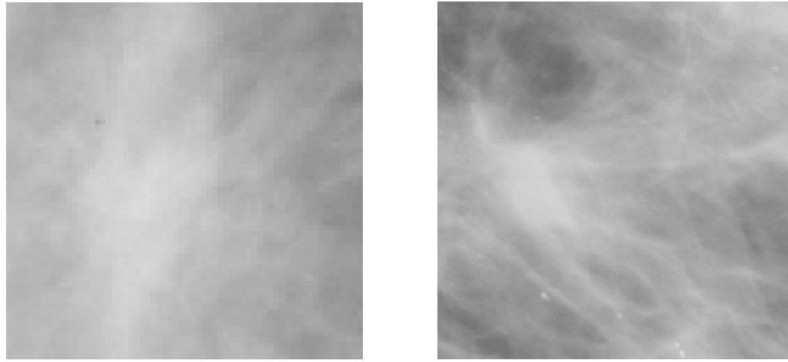


Fig. 2.7 Examples of Spiculated Mass

Round, oval and lobular shape: Masses in these categories imply a well defined smooth edge and is often benign. If their margin is not smooth, their shape alone does not tend to exclude malignancy. Mammograms with benign and malignant circumscribed masses are shown in fig. 2.8 and 2.9 respectively.

Irregular shape: irregular shapes tumor infiltrating edges are of more concern, because they imply indistinct and more often malignant cases.

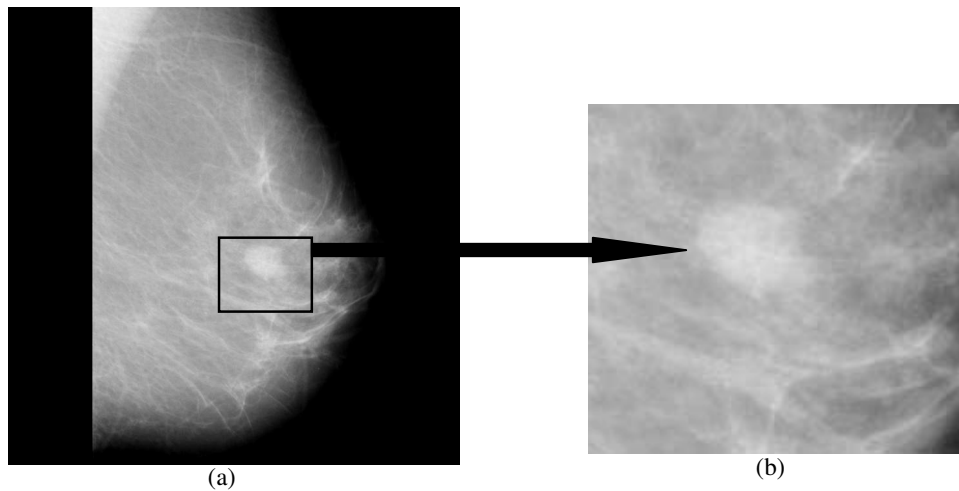


Fig. 2.8 Benign circumscribed mass (a) Original Mammogram with Benign circumscribed mass (b) Region containing Benign circumscribed mass

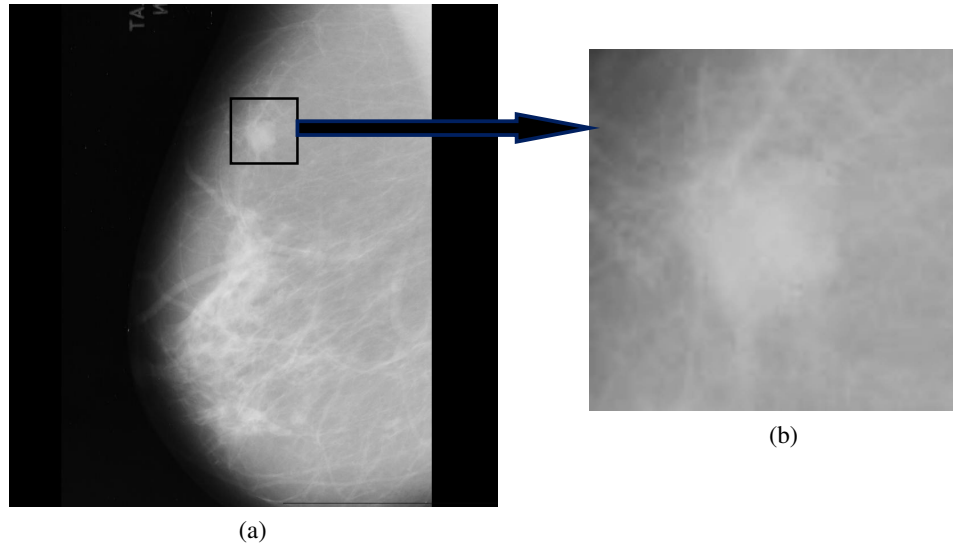


Fig. 2.9 Malignant Circumscribed mass (a) Original Mammogram with Malignant circumscribed mass (b) Region containing Malignant circumscribed mass

2.7.3 Architectural distortions

In this case, the normal outline of tissues is distorted, sometimes with no definable mass. It includes spiculations (line radiating from a center) and retraction of normal tissue planes where it causes contraction, which can be seen before an actual mass.

Architectural distortion occurs with healing after injury including previous biopsy and so it is critical to determine if the area has been injured. Benign causes of architectural distortion such as scarring tend to remain unchanged or improve. So one has to compare and observe if changes have really occurred. Fig. 2.10 illustrates the architectural asymmetry in mammograms.

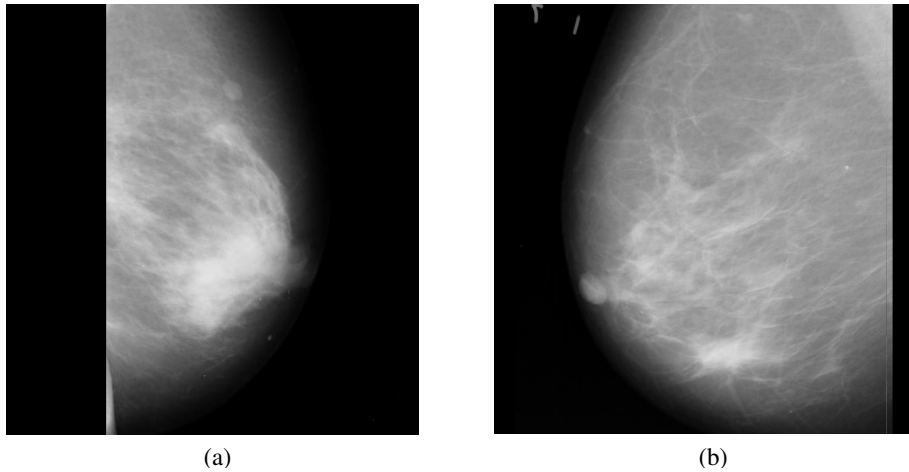


Fig 2.10 Examples of architectural asymmetry (a) Malignant (b) Benign

2.7.4 Spiculated lesions

Spiculated lesions appear as star shaped with blurred borders. They are almost always malignant. Generally, the lesion has a distinct central tumor mass with dense spicules radiating in all directions. The spicule length usually increases with tumor size. Occasionally translucent, oval or circular center or translucent areas within a loose structure and low density spicules characterize benign spiculated lesions. Examples of malignant and benign spiculated lesions are presented in fig 2.11

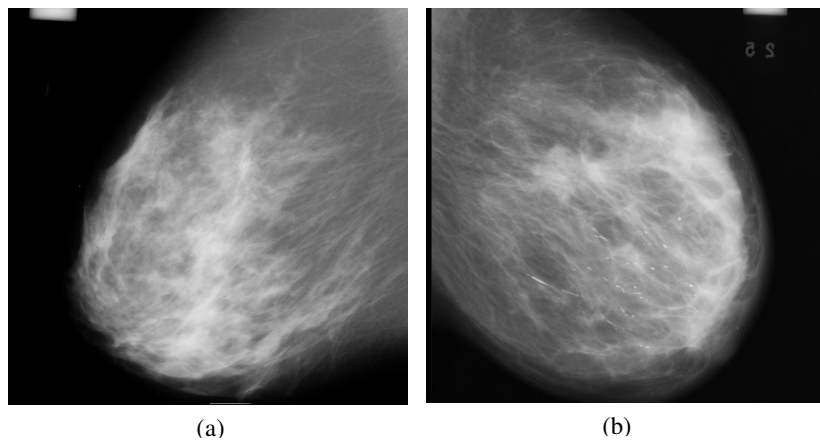


Fig 2.11 Examples of Spiculated Lesions (a) Benign (b) Malignant

2.8 Chapter Summary

This chapter presents a brief introduction to the breast cancer. Different breast imaging techniques that are available today are also discussed here. The most commonly used breast imaging modality is the mammography and is explained in detail. The abnormalities that can occur in the breast are described as well.

Chapter 3

A brief overview to the world of Fractals

This chapter provides an insight into the world of fractals used for the analysis of complex irregular structures. First, the basic properties of fractals are mentioned. The existence of fractals is explained by describing the space where fractals live. The mathematical foundations of fractals like the Iterated Function Systems (IFS) and Collage Theorem are given later. A succinct overview of the mathematical modeling of images using fractals is also presented.

3.1 From Euclidean geometry to Fractals

Conventional elementary mathematics classes deals with idealized shapes like triangles, circles, spheres, squares, etc. These form the basic building blocks of Euclidean geometry. But nature never exists in idealized shapes. Shapes such as coastlines, mountains and clouds are not easily described by traditional Euclidean geometry. Thus nature exhibits a higher degree or different level of complexity. The description of these shapes by Euclidean geometry is only an estimate of how closely they resemble to the Euclidean geometry. These irregular and fragmented patterns existing in nature lead to identifying a family of shapes called "fractals".

The word fractal was coined by Mandelbrot (Man 1982) in 1975 from the Latin adjective "*fractus*". The corresponding Latin verb "frangere" means "to break:" to create irregular fragments. Mandelbrot defined fractals as "A rough or fragmented geometric shape that can be subdivided in parts, each of which is (at least approximately) a reduced/size copy of the whole". He is often considered as the father of fractal geometry.

But, the works on fractals were started long back before Mandelbrot. Their description began with classical mathematics and mathematicians like Georg Cantor (1872), Giuseppe Peano (1890), David Hilbert (1891), Helge von Koch (1904), Waclaw Sierpinski (1916) Gaston Julia (1918) and Felix Hausdorff (1919) were the pioneers in this field (Pei 1991).

Methods of classical geometry and calculus are unable for studying fractals and alternate techniques are required (Fal 2003). Thus, fractal geometry developed by Mandelbrot, forms a new branch of mathematics, which is appropriate for the irregular shapes in the real world.

These definitions lead to the three important properties of fractals which are given below.

3.2 Properties of Fractals

The three basic properties of fractals are

- a. Self similarity
- b. Fractal dimension
- c. Formation by Iteration

3.2.1 Self Similarity

Self similarity stands for "similar to itself". An example of a natural structure with self similarity is a cauliflower. The cauliflower contains branches or parts which when removed and compared with the whole, are very much the same, only smaller. These clusters again can be decomposed into smaller clusters, which again look similar to the whole as well as to the first generation branches. This self similarity carries forward, for about three or four stages. After that the structure is too small for further dissection. The self similarity of cauliflower is shown in fig 3.1.

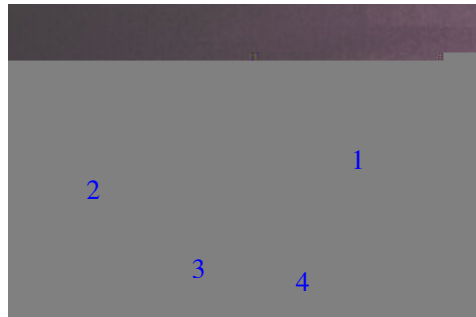


Fig. 3.1 Four pieces of the same single cauliflower is shown in 1, 2, 3 and 4.






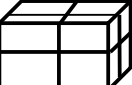
In a mathematical idealization, the self similarity property of a fractal may be continued through infinite number of stages (Pei 1991). But all self similar structures are not fractals. For example, a line segment or a square or a cube, can be broken into small copies which are obtained by similarity transformation. But, these structures are not fractals. There comes the second property of fractals: the fractal dimension.

3.2.2 Fractal Dimension

The fractal dimension is a statistical quantity that gives an indication of how completely a fractal appears to fill space, as one zooms down to finer and finer scales. Before coming to fractal geometry, the dimensions in the Euclidean geometry are considered. A point has no dimension as it does not have length, width or height. A straight line has dimension 1 since it has finite length, but no width or height. But a plane has two dimensions length and width but no height. For a cube, the dimension is three, as it has three dimensions length, width and depth extending to infinity in all three directions.

For further investigations on dimension, take a self-similar figure like a line segment, and double its length. Doubling the length gives two copies of the original segment. Considering another self similar figure like a square, when the length and

Table 3.1 Relation between Scaling factor, Number of copies and Dimension

Figure	Dimension (D)	Scaling Factor (r)	Number of copies (N)
Point .	0	-	-
Line 	1	2	 $2=2^1$
Plane 	2	2	 $4=2^2$
Cube 	3	2	 $8=2^3$

width are scaled by two, it gives four copies of the original cube. Similarly eight copies are obtained when a cube is doubled in its length, width, and height. The table

3.1 shows the relation between the number of copies (N), scaling factor (r) and the dimension (D). Thus the number of copies N is given by:

$$N = r^D \quad (3.1)$$

Taking logarithm on both sides,

$$\log(N) = D \log(r)$$

or
$$D = \frac{\log(N)}{\log(r)} \quad (3.2)$$

Thus, the dimension is an integer in Euclidean geometry. Now, consider the most common example of fractal, the Sierpinski triangle, introduced by the Polish mathematician Waclaw Sierpinski (1882-1969) in 1916 (Pei 1991). The basic construction of the Sierpinski triangle (gasket) is as follows.

Draw an equilateral triangle in a plane. Pick the midpoints of its three sides (fig 3.2a). When the three midpoints are joined four triangles are formed as in fig. 3.2(b), out of which the central one is dropped. This is the basic construction step.

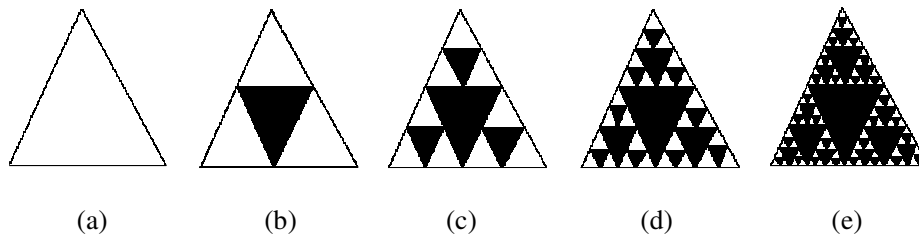


Fig 3.2 Basic Construction Steps of Sierpinski triangle

Thus, after the first step, three congruent triangles, which are exactly half the size of the original triangle, which touch at three points i.e. common vertices of the two contiguous triangles. The same procedure is followed with the three remaining triangles and repeat the basic step as often as desired. Therefore, starting with one triangle, succeeding stages produces, 3, 9, 27, 81, 243.....triangles, each of which is an exact scaled down version of the triangles in the preceding step.

Now, defining the dimension of Sierpinski triangle, when the length of the sides is doubled (since the black triangles are holes, they are not counted), three

copies of the original triangles are obtained. Following the convention in equation (3.1), number of copies $3 = 2^D$, where D is the dimension. From equation (3.2)

$$D = \frac{\log(3)}{\log(2)} = 1.5849 \quad (3.3)$$

Thus the fractal dimension is not an integer. So fractals are geometrical objects that have a non integer or fractional dimension.

3.2.3 Formation by Iteration

Fractals are often formed by an iterative procedure. To make a fractal, take a familiar geometric figure, like a triangle, a line segment, etc and operate on it so that the new figure is more "complicated" in a special way. Then, in the same way, operate on that resulting figure, and get an even more complicated figure. Repeat the process again and again...and again. It should be repeated many times. Consider the previous example of Sierpinski triangle, the final figure obtained is very complex compared to the starting figure of a triangle.

Thus, a fractal is a design of infinite details. It is created using a mathematical formula. No matter how closely they are looked at, a fractal never loses its detail. It is infinitely detailed, yet it can be contained in a finite space. Thus, fractals are generally self similar and are independent of scale.

Some of the popular fractal sets defined by mathematicians, other than the Sierpinski triangle, are Cantor set, Koch curve, Julia set, Mandelbrot set etc. and are shown in the figs.3.3(a)-3.3(d).

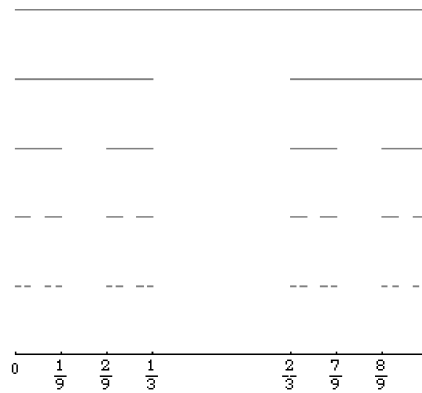


Fig 3.3 (a) Basic steps for the construction of Cantor set

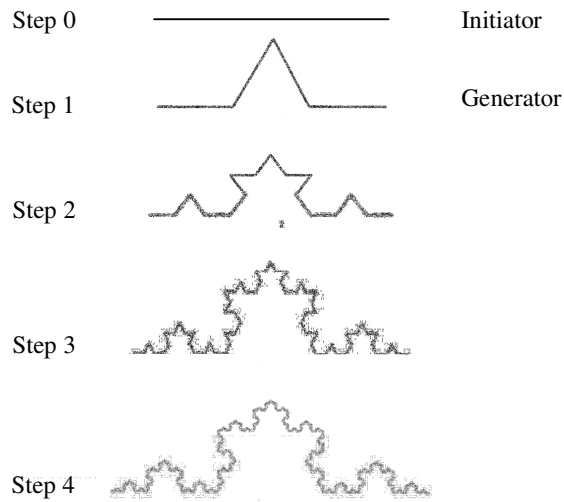


Fig 3.3 (b) Basic steps for the construction of Koch Curve

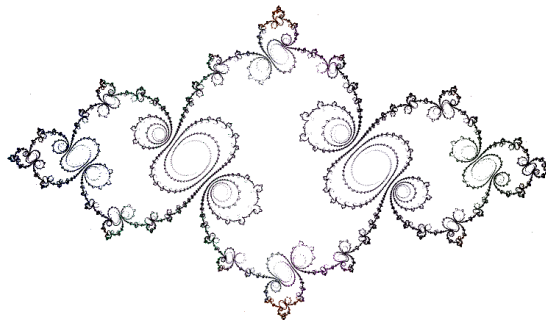


Fig 3.3 (c) Julia Set

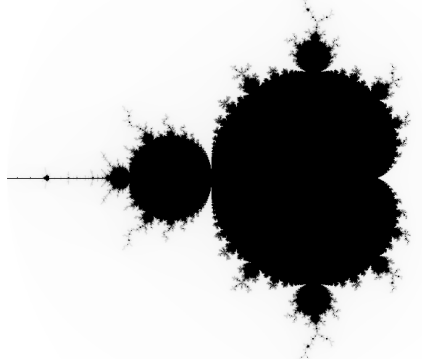


Fig.3.3 (d) Mandelbrot Set

The mathematical foundations of fractals can be explained by considering the analogy of a simple Xerox machine with feedback incorporated in it and is explained in the next section.

3.3 Multiple Reduction Copy Machine

Consider a simple Xerox (copy) machine that takes an image as the input. It has several independent lens systems, each of which reduces the input image and places it somewhere in the output image. The assembly of all reduced copies in some pattern is finally produced as output (Pei 1991). Schematic of such a copy machine is given in fig. 3.4.

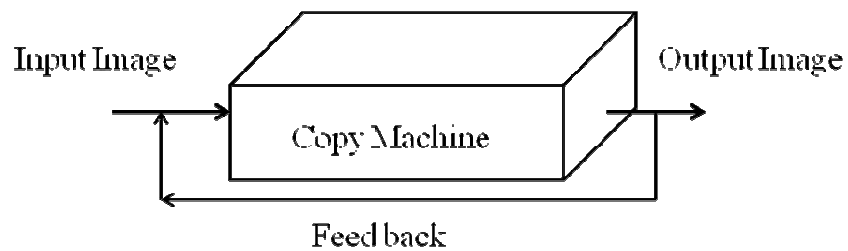


Fig. 3.4 Schematic of Multiple reduction copy machine

The crucial point in the design is that the machine runs in a feedback loop; its own output is fed back as its new input again and again. When there are more than one reduction lens, powerful and exciting results are obtained. Since the output image is the reduced copy of the original image, that copy is *similar* to the original. The process to generate a copy is called *similarity transformation* or *similitude*. After ten or more cycles, any initial image will reduce to just a point. Such a machine is referred to as the Multiple Reduction Copy Machine (MRCM).

Now; consider the MRCM with three lens systems, each of which is set to reduce the input image by a factor of $1/2$. The resulting copies are assembled in the configuration of an equilateral triangle. Fig.3.5 shows an MRCM with a square as the input image and the reduced copies of the square are placed in the form of an

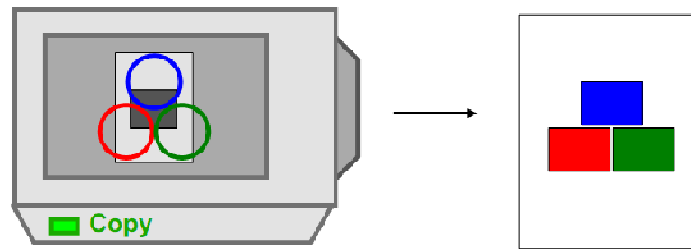


Fig. 3.5 Multiple Reduction Copy Machine with three reduction lenses and places the input figure in the form of an equilateral triangle

equilateral triangle. Fig. 3.6 shows the effect of the machine run three times, beginning with different initial images. Looking into the output images obtained in fig. 3.6, it is observed that the same final structure is approximated as the machine is run infinite number of times. It doesn't matter in the least, whether the input images to the machine are rectangles, triangles, or any other shape, the same final image is approached in each case - the Sierpinski gasket. This final image is totally independent of the initial image with which the operation was started. Thus, in mathematical terms, it is a process which produces a sequence of results tending towards one final object which is independent of how the process began. This property is called *stability*. The final image to which the algorithm converges is called the *attractor*. Moreover, when the machine is started with the attractor, then nothing

happens, or the attractor is left invariant or fixed. The arrangement of the lens and its reduction factor determines the final image it produces. Thus it is equivalent to applying certain transformations on the input image, which are called the *affine transformations*.

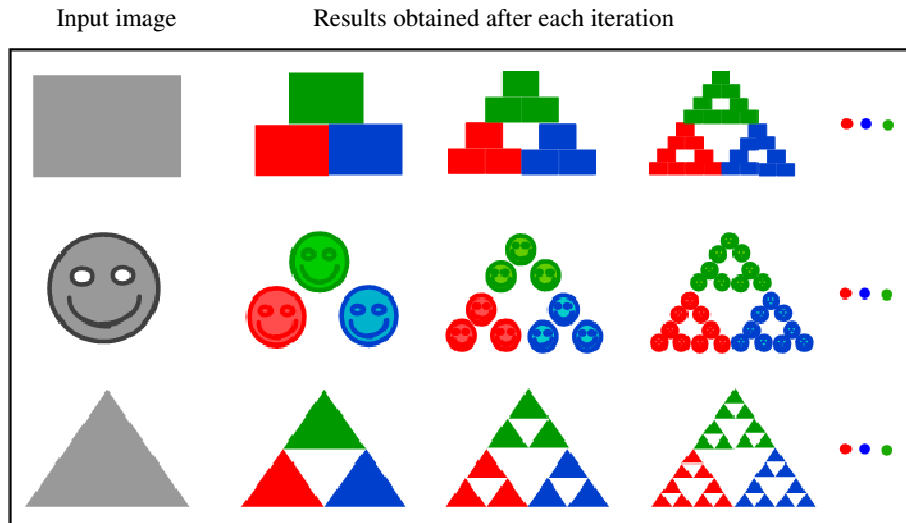


Fig. 3.6 The input images to the MRCM and their corresponding output images obtained after three iterations

The mathematical foundations of the fractals can be approximated to the working of the above mentioned Multiple Reduction Copy Machine (MRCM) algorithm.

3.4 Mathematical Foundations

Irregular sets such as those seen in section 3.2 provide a much better representation of many natural phenomena than the figures of classical geometry. Fractal geometry provides a general frame work for the study of such irregular sets. A few basic definitions in the fractal space are explained below (Fal 2003).

3.4.1 The Space of Fractals

Fractal geometry is concerned with the structure of sub-sets of various very simple “geometrical” spaces (Bar 1988). Such a space is denoted by \mathbf{X} . Fractal is just a sub-set of a space. Though the space is simple, the fractal sub-set is geometrically complicated.

The space \mathbf{X} is a set. The points of the space are the elements of the set. The nomenclature “space” implies that there is some structure to the set, some sense of which points are close to which. Example: when $\mathbf{X} = \mathfrak{R}$, \mathfrak{R} denotes the set of real numbers. Each “point” $x \in \mathbf{X}$ is a real number. When $\mathbf{X} = \mathfrak{R}^2$, it is the Euclidean plane or the coordinate plane of calculus. Any pair of real numbers $x_1, x_2 \in \mathfrak{R}$ determine a single point in \mathfrak{R}^2 . That is, in \mathfrak{R}^2 space, any point is a function of two co-ordinates (x, y) . Any image can be considered as a set in the \mathfrak{R}^2 space or in other words images are all possible subsets of \mathfrak{R}^2 space.

The next section gives certain definitions which are necessary for building the mathematical foundations of fractals.

3.4.1.1 Metric spaces

A metric space (\mathbf{X}, d) is a space \mathbf{X} together with a real valued function $d : X \times X \rightarrow \mathfrak{R}$, which measures the distance between pairs of points x and y in \mathbf{X} . It is required that d should obey the following axioms.

- (1) $d(x, y) = d(y, x) \quad \forall x, y \in X$
- (2) $0 < d(x, y) < \infty \quad \forall x, y \in X, x \neq y$
- (3) $d(x, x) = 0 \quad \forall x \in X$
- (4) $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$

The fourth axiom is called the *triangular inequality*.

Such a function is called a *metric*. The concept of shortest paths between points in a space, *geodesics*, is dependent on the metric. The metric may determine a geodesic

structure of the space. Geodesics on a sphere are great circles and in the plane with the Euclidean metric they are straight lines.

Fractal geometry is concerned with the description, classification, analysis and observation of subsets of metric spaces (\mathbf{X}, d) . The metric spaces are usually of an inherently simple geometric character and the subsets are typically geometrically complicated. There are a number of general properties for the subsets of metric spaces, which occur over and over again, which are very basic, and which form part of the vocabulary for describing fractal sets and other subsets of metric spaces.

3.4.1.2 Cauchy Sequence

A sequence $\{x_n\}_{n=1}^{\infty}$ of points in a metric space (\mathbf{X}, d) is called a Cauchy sequence if, for any given number $\varepsilon > 0$, there is an integer $N > 0$, so that

$$d(x_n, x_m) < \varepsilon, \forall m, n > N \quad (3.4)$$

The above equation implies that, further along the sequence if one goes, the closer together become the points in the sequence as in fig.3.7. Just as the points move close together as one move along the sequence, it will finally converge to a point.

A sequence $\{x_n\}_{n=1}^{\infty}$ of points in a metric space (\mathbf{X}, d) is said to converge to a point

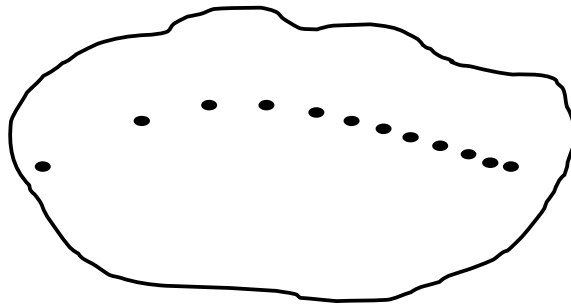


Fig.3.7 Cauchy Sequence: Points gets closer and closer along the sequence

$x \in X$ if, for any given number $\varepsilon > 0$, there is an integer $N > 0$ so that

$$d(x_n, x) < \varepsilon, \quad \forall n > N \quad (3.5)$$

In this case the point $x \in X$, to which the sequence converges, is called the limit point of the sequence;

$$x = \lim_{n \rightarrow \infty} x_n \quad (3.6)$$

If a sequence of points $\{x_n\}_{n=1}^{\infty}$ in a metric space (X, d) converges to a point $x \in X$, then $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence.

A metric space (X, d) is complete if every Cauchy sequence $\{x_n\}_{n=1}^{\infty}$ in X has a limit point $x \in X$.

3.4.2 Affine Transformations

Consider mappings from \mathfrak{R}^2 to \mathfrak{R}^2 such that the new points formed are members of the space \mathfrak{R}^2 . Let the mapping be given by, $(x, y) \rightarrow (x', y')$ or in other words, the new (x', y') formed are functions of previous points x and y . i.e

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = f \begin{pmatrix} x_n \\ y_n \end{pmatrix} \quad (3.7)$$

A transformation $\omega: \mathfrak{R}^2 \rightarrow \mathfrak{R}^2$, of the form

$$(x_{n+1}, y_{n+1}) = (ax_n + by_n + e, cx_n + dy_n + f) \quad (3.8)$$

where a, b, c, d, e and f are real numbers is called affine transformation.

or

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \quad (3.9)$$

$$= Ax + t$$

Here $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is two dimensional, 2×2 real matrix and t is the column

vector $\begin{pmatrix} e \\ f \end{pmatrix}$, which cannot be distinguished from the co-ordinate pair $(e, f) \in \mathfrak{R}^2$. The

matrix A can always be written in the form:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} r_1 \cos \theta_1 & -r_2 \sin \theta_2 \\ r_1 \sin \theta_1 & r_2 \cos \theta_2 \end{pmatrix} \quad (3.10)$$

where (r_1, θ_1) are the polar coordinates of the point (a, c) and $(r_2, (\theta_2 + \pi/2))$ the polar coordinates of the point (b, d) . The linear transformation in \mathfrak{R}^2 given by equation 3.11, maps any

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow A \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.11)$$

parallelogram with a vertex at the origin to another parallelogram with a vertex at the origin. The parallelogram is turned over by the transformation. A general affine transformation $\omega(x) = Ax + t$ in \mathfrak{R}^2 consists of a linear transformation, A, which deforms space relative to the origin, i.e., flipping, rotation etc, followed by a translation or shift specified by the vector t as shown in fig. 3.8. Let $\omega : X \rightarrow X$ be a transformation on a metric space. A point $x_f \in X$ such that $f(x_f) = x_f$ is called a fixed point of the transformation. The fixed points of a transformation are very important. They are pinned into space i.e. they are not changed by the transformation. They restrict the motion of the space under non violent, nonripping transformation of bounded deformation.



Fig 3.8 Deformation obtained for the parallelogram

3.4.3 The Contraction Mapping Theorem

Let x and y be two points in space with a distance $d(x, y)$ between them. Define two functions $f(x)$ and $f(y)$. They will produce two new different points with a distance

of $d(\omega(x), \omega(y))$ between them. A transformation $\omega : \mathbf{X} \rightarrow \mathbf{X}$ on a metric space is called contractive or contraction mapping if there is a constant $0 \leq s < 1$ such that

$$d(\omega(x), \omega(y)) \leq s \cdot d(x, y) \quad \forall x, y \in X \quad (3.12)$$

where s is called the contractivity factor. For the contractivity factor to be less than one the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix} < 1$.

A contraction mapping has at most one fixed point. Moreover, the Banach fixed point theorem states that every contraction mapping on a nonempty complete metric space has a unique fixed point, and that for any x in X the iterated function sequence converges to the fixed point.

3.4.3.1. Banach's Contraction Mapping Theorem.

Let (X, d) be a complete metric space, and $\omega(x)$ be the transformation applied on x , which is contractive mapping. Then

$$d(\omega(x), \omega(y)) \leq \lambda d(x, y), \quad 0 < \lambda < 1, \quad \forall x, y \in X \quad (3.13)$$

Then ω has a unique fixed point. When there are a set of transformations, $\omega_1, \omega_2, \omega_3$ with contractivity factors s_1, s_2 and s_3 respectively, then the contractivity of $\omega = \omega_1 \cup \omega_2 \cup \omega_3$ is $s = \max\{s_1, s_2, s_3\}$.

3.4.4. Iterated Function System (IFS)

An Iterated Function System (IFS) consists of a complete metric space (\mathbf{X}, d) together with a finite set of contraction mappings $\omega_n : X \rightarrow X$, with respective contractivity factors, s_n , for $n=1,2,\dots,N$. Let a transformation defined as $\Omega : H(X) \rightarrow H(X)$ by the transformation Ω be applied on the set B in the complete metric space.

$$\begin{aligned} \Omega(B) &= \omega_1(B) \cup \omega_2(B) \cup \dots \cup \omega_n(B) \\ &= \bigcup_{n=1}^n \omega_n(B), \quad \forall B \in H(X) \end{aligned} \quad (3.14)$$

Ω is a contraction mapping with contractivity factor. Then $h(\Omega(B), \Omega(C)) \leq s.h(B, C)$, $\forall B, C \in H(X)$. Its unique fixed point $A \in H(X)$ obeys

$$A = \Omega(A) = \bigcup_{n=1}^n \omega_n(A) \quad (3.15)$$

and is given by $A = \lim_{n \rightarrow \infty} \Omega^{0n}(B)$ for any $B \in H(X)$. The fixed point $A \in H(X)$ is called the attractor of the IFS. This implies that when starting with any set or image, on the repeated application of the IFS, it will ultimately converge to the fixed point or another fixed image. This is similar to the property of MRCM discussed in section 3.3.

Till this point, the problem of how to generate the image when a set of IFS is given is considered. But, an inverse problem exists, which is, given a fractal object how to generate the IFS. This is based on the collage theorem which is given in the next section.

3.4.5. Collage Theorem

This theorem is the corollary of the Banach's fixed point theorem and the central theorem with which the inverse theorem is based. The explanation given here is based on the derivation by Barnsley (Bar 1988). According to the Banach's fixed point theorem the distance between the two point's x_m and x_n in the sequence is given by:

$$d(x_m, x_n) \leq d(x_0, x_1) \frac{\lambda^m}{1 - \lambda} \quad (3.16)$$

Setting $m=0$ and $n=\infty$ the above equation becomes:

$$d(x_0, x_\infty) \leq d(x_0, x_1) \frac{1}{1 - \lambda} \quad (3.17)$$

Let (X, d) be a complete metric space. Let L , a fractal object, $L \in H(X)$, be given and let $\varepsilon \geq 0$ be also given. Choose an IFS such that $\{X; (\omega_0), \omega_1, \omega_2, \dots, \omega_n\}$ with contractivity factor $0 \leq s < 1$, so that

$$h\left(L, \bigcup_{n=1}^n \omega_n(L)\right) \leq \varepsilon \quad (3.18)$$

where $h(d)$ is the Hausdorff metric. For a complete metric space, the Hausdorff distance between points A and B is defined by

$$h(A, B) = d(A, B) \vee d(B, A) \quad (3.19)$$

The notation \vee is used to mean the maximum of the two real numbers. Then

$$h(L, A) \leq \frac{\varepsilon}{1-s} \quad (3.20)$$

where A is the attractor of the IFS. Or Equivalently:

$$h(L, A) \leq \frac{1}{1-s} h\left(L, \bigcup_{n=1}^n \omega_n(L)\right) \quad \forall L \in H(X) \quad (3.21)$$

The theorem tells that in order to find an IFS whose attractor is “close to” a given set, find a set of transformations, that is, contraction mappings on a suitable space within which the given set lies, such that the union, or collage of images of the given set under the transformations is near to the given set. Nearness is measured using Hausdorff metric.

Thus while fractal image coding, find out which transformation when applied to the whole image, will give a part of the whole image. Then make collage, and this collage should be as close as the whole image.

3.5 Fractal Image Coding

The theorems mentioned above in section 3.4, which employs the self similarity property of the fractal images are exploited for fractal image coding. Since deterministic fractal objects are redundant objects, in the sense that, they are made up of transformed copies of either themselves or parts of themselves. This property is mainly used in compression of images based on fractal coding. For a general image, there is no need to store all the parts of the image, since affine redundancies will be present in parts of the image. Image redundancy can be exploited by modeling it, as is present in fractal objects.

In the basic method of fractal image coding proposed by Barnsley, the image to be coded is divided into blocks, called range blocks as shown in fig.3.9. Then the entire image is searched for a corresponding domain block such that, the best block when coded gives the range block. The size of the domain is generally chosen to be greater than range, to ensure contractive mapping between domain and range. Usually the size of the domain is chosen to be twice that of the range block. The domains are scaled and rotated in different directions and compared for best mapping between the domain and the range and thus finding the affine transformation, mapping the domain and the range block. This is computationally intensive as the entire image has to be searched to find the best mapping from a domain to the range. Then the domain and range locations along with the six parameters (a , b , c , d , e , and f) computed using equation 3.9, are stored. Finally, collage theorem is used to get the coded image from these parameters. Thus instead of storing the image block its corresponding fractal codes are stored. This tremendously reduces the storage space leading to a large compression ratio.

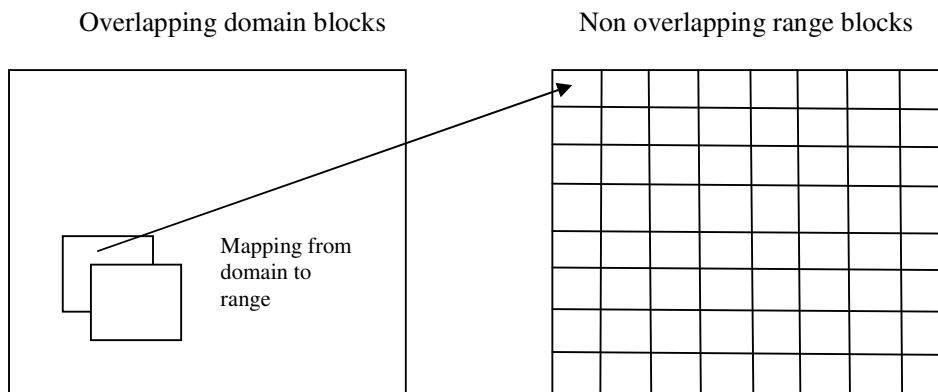


Fig 3.9 Image divided into non overlapping range blocks and the most suitable domain block is found by searching the entire image

The above method introduced by Barnsley has been modified by many mathematicians to reduce the time required to find the matching domain for a range

block. This fractal block coding forms the basis for modeling of mammograms used to identify the presence of microcalcifications described in chapter 5.

3.6 Literature Survey

Fractal objects which are generated from the mathematical theory of Iterated Sequences, were first “tagged” as mathematical “curiosities” or “monsters” by mathematicians in the beginning of the twentieth century.

Mathematicians Gaston Julia, Karl Weierstrass and Waclaw Sierpinski were among the first to explain the geometric properties of fractal (Fal 2003). They lacked the tools to properly analyze and understand them. They remained in nearly complete oblivion till they were rediscovered in the 1970’s. It began by the pioneering work of Benoit Mandelbrot who also coined the name fractals (Man 1982).

Benoit Mandelbrot, a French mathematician, is known as the father of fractal geometry. He was the first to confer the term fractal, and described these structures and his ideas in his book ‘The Fractal Geometry of Nature’ (Man 1982). His most famous contribution was a set of points in the complex plane, now known as the Mandelbrot set, which form an amazing fractal. One of the most impressive things about the development of this fractal, as with all early fractals, is that it was drawn by hand rather than generated with computing technology. Since this pioneering in fractal history, the Mandelbrot set has been drawn and redrawn time and time again, not for the sake of complex dynamics, but for the creation of art.

Due to availability of computers and automatic graphic tools, it was possible to render and visualize them as complex, beautiful, often realistic looking objects or scenes. Later on fractals have been a part of a set of tools in a variety of fields in physics, where they are closely related to chaos theory.

Fractals have blossomed tremendously in the past few years and have helped to reconnect pure mathematics research with both the natural sciences and computing. Fractal geometry and its concepts has become the central tool in most of the natural sciences: physics, chemistry, biology, geology, meteorology and materials science. There are lots of books available in literature describing the basics of fractals (Pei

1988, Hel 2007, Tri 1997). Selection of the classical mathematical papers by masters like, G.Cantor, Sierpinski, B. Mandelbrot, Felix Hausdorff related to fractal geometry is compiled (Edg 2004). Wavelets and fractals were analyzed based on probability by Jorgensen (Jor 2006).

Barnsley (Bar 1988) explains the method of fractal image compression. The original approach of fractal image coding devised by Barnsley, became a practical reality when Jacquin (Jaq 1992, Jaq 1993) introduced the partitioned iterated function system in that each of the individual mappings operates on a subset of the image, rather than the entire image. Franceschetti and Riccio (Fra 2007) mention the fundamentals of stochastic fractals. The chaos theory was detailed by Stupnicka in (Stu 2003) and Schroeder (Sch 1991).

Iftekharruddin *et. al* (Ift 2003) implemented an improved Piecewise Modified Box Counting (PMBC) and Piecewise Triangular Prism Surface Area (PTPSA) methods, to find the differences in intensity histogram and fractal dimension between normal and tumor images to detect and locate the tumor in the brain MR images.

A method incorporating gray relational pattern analysis into the self-organizing feature maps (SOFM) network to develop a GSOFM network to speed up the encoding time to about 500 was proposed by Jianwei and Jinguang (Jnw 2008).

Berizzi *et. al* (Ber 1997) proposed a two-dimensional fractal model of the sea surface, by means of the solution of the sea hydrodynamic differential equations and based on the band-limited Weierstrass-Mandelbrot (WM) fractal functions. Fractal geometry is used to take into account of the multi scale nature of the sea and to give a better description of the fine structure of the sea surface and they have also determined a directional wave spectrum of the sea.

Kinsner *et.al.* (Kin 2009) describes a novel approach of fractal modeling and coding of residuals for excitation in the linear predictive coding of speech. The new speech coder implemented using the piecewise self-affine fractal model gave a signal-to-noise ratio of 10.9 dB.

A new type of turbulence model based on fractals, applicable both in a Reynolds averaged Navier–Stokes (RANS) and a large-eddy simulation (LES) formulation which assumes an isotropic behavior for the turbulent viscosity was

developed by Giacomazzi *et.al* (Gia 1999). This can be applied for simulating turbulent combustion irrelevant of its mode (premixed or non-premixed) and was able to turn itself off in the laminar zones of the flow, and in particular near walls.

Partial discharge (PD) occurring in XLPE power cables is a cryptic phenomenon with detectable features differing in a thousand ways. The authors have investigated the use of fractal features for recognition of 3-d PD patterns as a fractal surface and it was found that two fractal features, fractal dimension and self-similar characteristic, possess reasonably good pattern recognition (Luo 2002). The study of the nature of the fractal features of 3-d PD patterns provided an efficient method for recognizing and picking-up the faint PD pulse from noise based on fractal theory.

Zho *et.al* (Zho 2010) proposed a multirange fractal model to calculate transition curves of multirange fractals by utilizing relevant fractal dimensions of Sn melt at different temperatures.

Most natural images, such as geographical images, are all textural in nature. In remote sensing images, different regions possess different texture and have different multifractal exponents. These properties were utilized for image segmentation by Du and Yeo (Du 2002).

In astrophysics, structures with fractal characteristics are important. A novel approach to characterize prefactors of cover functions, like lacunarity, based on the formalism of regular variation (in the sense of Karamata) is proposed by Stern (Ste 1997). This approach allowed in deriving bounds on convergence rates for scaling exponent algorithms and provided a more precise characterization for fractal-like objects of interest for astrophysics.

Deering and West (Dee 1992) related the complexity of physiological structures like cochlea and lungs to fractal geometry.

A watermarking method which utilizes a special type of orthogonalization fractal coding method was proposed by Pi *et. al* (PiH 2006). Here, the watermark embedding procedure inserts a permuted pseudo-random binary sequence into the quantized range block means and was found to be robust against common signal and

geometric distortion such as JPEG compression, low-pass filtering, rescaling, and clipping.

Winding currents in transformers vary to different extent depending upon the type of impulse fault. Fractal analyses using features like fractal dimension and lacunarity, of such complex current waveforms have been reported by Purkait and Chakravorti (Pur 2003) for classification of impulse faults in transformers. Experimental results obtained for a 3 MVA transformer and simulation results obtained for 3 MVA, 5 MVA and 7 MVA transformers are presented.

The application of fractals in stock markets is explained by Mandelbrot (Man 1997). Losa *et. al* (Los 2005) presented the biological and medical applications of fractals.

The fractal nature of the imaging modalities that measure flow vector fields (flow-sensitive MRI and Doppler ultrasound) were modeled by Tafti *et. al* (Taf 2010) using FBM along with vector models then used them to analyze 3D flow measurements obtained using phase-contrast MRI.

Thus, it can be seen that fractals find applications in a various fields. In the present research, the fractal dimension is used for the classification of mammograms while the self similarity property is used for fractal modeling of mammograms.

3.7 Chapter Summary

The new word fractal was introduced by Mandelbrot to encompass all the complex geometric shapes. Unlike Euclidean shapes, fractals have dimension which is non integer or fractional. This chapter provides a brief introduction to fractals. The basic concept and the theoretical background of fractals are presented. Literature survey of the application of fractals in diverse fields is also incorporated here.

Chapter 4

Development of new Fractal features for the Classification of Mammograms into Normal, Benign and Malignant

Abnormalities in the mammograms include masses and microcalcifications, which can be benign or malignant. Due to the presence of these abnormalities, the regularity of the mammogram structure is altered, which changes its fractal dimension. This chapter deals with the classification of mammograms based on fractal dimension and features. To compute fractal dimension, three methods viz. the Differential Box Counting Method, Blanket Method and Triangular Prism Surface Area method were used. Since it was observed that fractal dimension cannot uniquely distinguish between different classes of mammograms, six different fractal features were derived from the above mentioned fractal dimension estimation methods. The classification performances of these classifiers are evaluated using Receiver Operating Characteristics (ROC).

Chapter 4

Development of new Fractal features for the Classification of Mammograms into Normal, Benign and Malignant

Abnormalities in the mammograms include masses and microcalcifications, which can be benign or malignant. Due to the presence of these abnormalities, the regularity of the mammogram structure is altered, which changes its fractal dimension. This chapter deals with the classification of mammograms based on fractal dimension and features. To compute fractal dimension, three methods viz. the Differential Box Counting Method, Blanket Method and Triangular Prism Surface Area method were used. Since it was observed that fractal dimension cannot uniquely distinguish between different classes of mammograms, six different fractal features were derived from the above mentioned fractal dimension estimation methods. The classification performances of these classifiers are evaluated using Receiver Operating Characteristics (ROC).

4.1 Introduction

It is always desirable to develop computer-based methods to distinguish between benign masses and malignant tumors while considering the traumatic nature and cost of biopsy. Such methods can help in performing initial screening or second reading of mammograms, and lend objective tools to help radiologists in analyzing difficult cases and decide on biopsy recommendations.

The goal of this chapter is to develop an efficient new fractal feature derived from fractal dimension, to assist radiologists' in categorizing mammograms into normal and abnormal.

Mandelbrot defined a number, associated with each fractal, called its fractal dimension. It reflects the measure of complexity of a surface, the scaling properties of the fractal i.e. how its structure changes when it is magnified. Thus fractal dimension gives a measure of the irregularity of a structure.

Abnormal masses and microcalcifications such as benign and malignant are used in this research. The fig 4.1 shows the different classes of mammograms. Measures that can quantitatively represent shape, texture and complexity can assist in the classification of mammograms into benign and malignant.

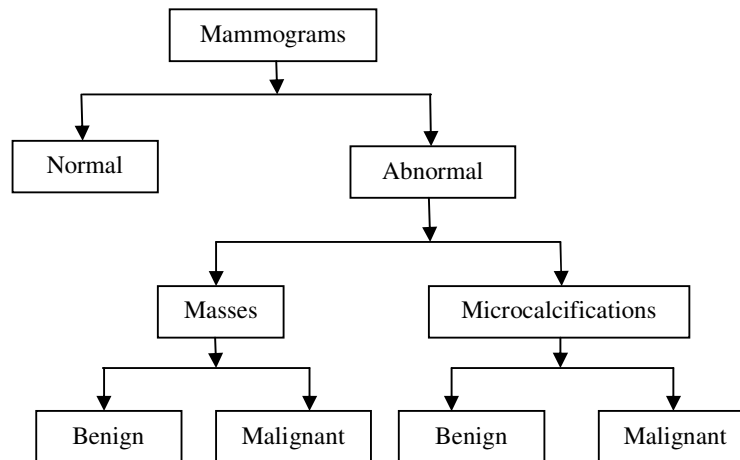
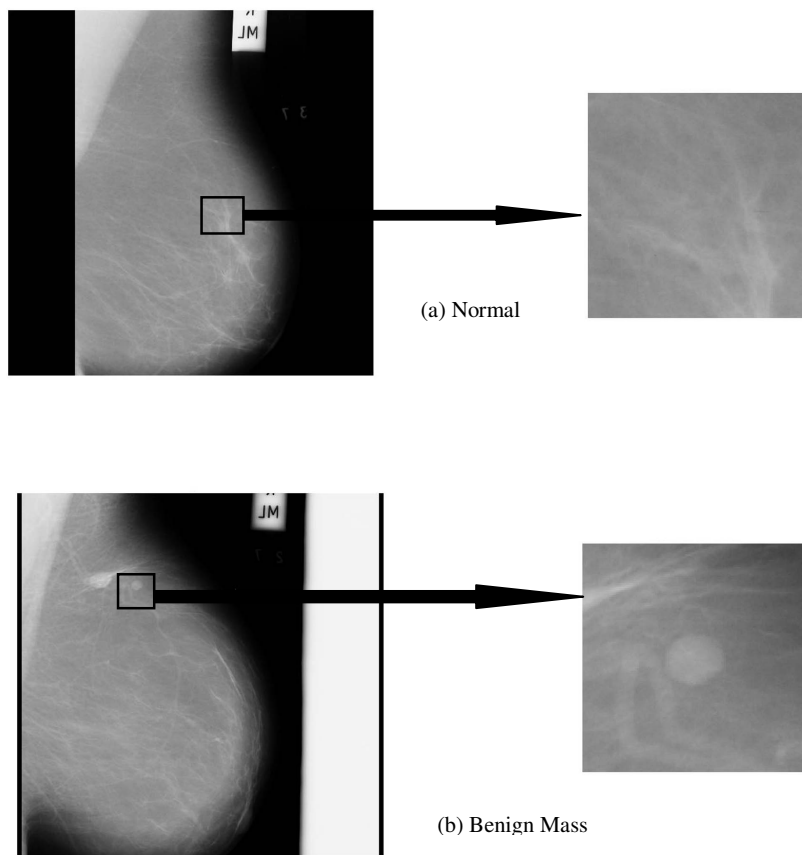


Fig 4.1 Different classes of mammograms

Normal mammograms usually have a regular structure, but due to the presence of the abnormal tissues the complexity of abnormal mammograms increases. Thus, naturally they will have higher fractal dimension. Malignant tumors are generally rough and have more irregularity whereas benign masses commonly have smooth, round, oval contours.

Fig. 4.2 shows examples of mammograms from each class. In the normal mammogram the ducts and tissues patterns are clearly visible. This makes the interpretation of mammograms difficult, if microcalcifications or masses are embedded in it.



Contd....

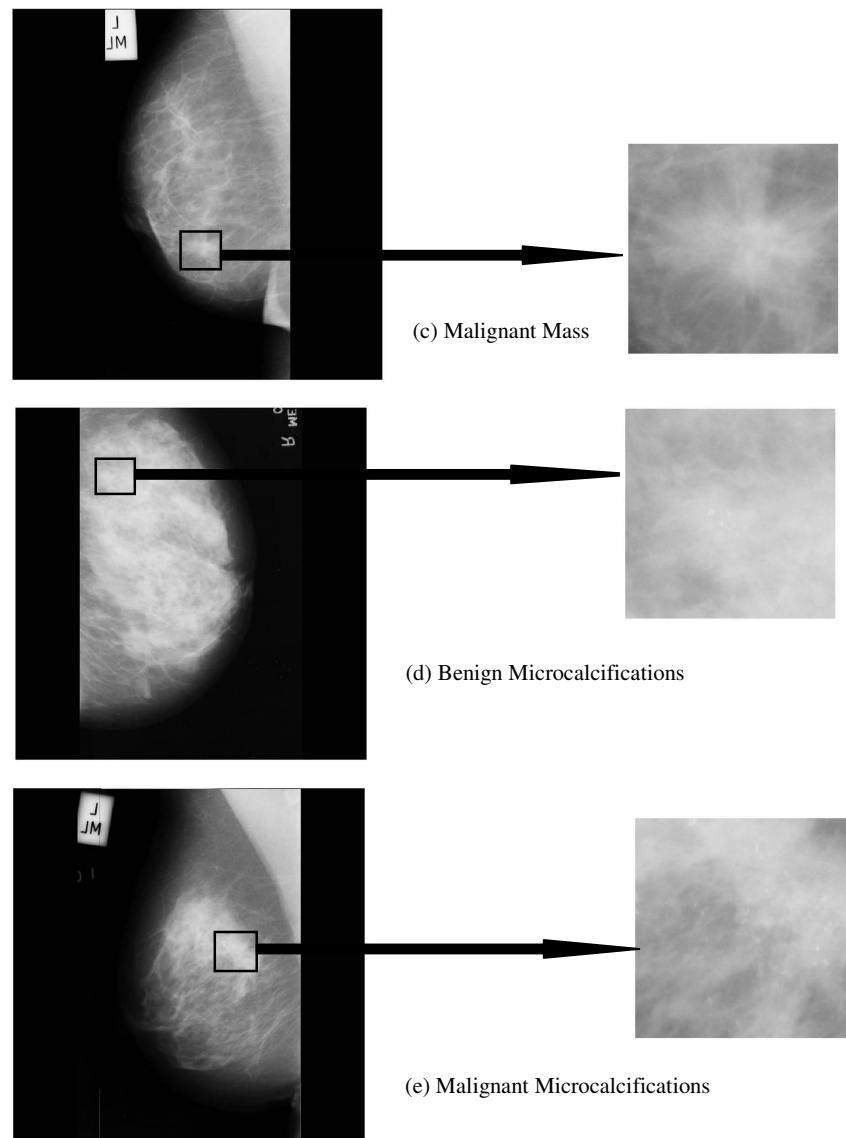


Fig.4.2 Different Classes of Mammograms: Original and ROI taken out from the mammogram (a) Normal (b) Benign Mass (c) Malignant Mass (d) Benign Microcalcifications (e) Malignant Microcalcifications

In this research, the property of fractional Dimension (FD) is used for the classification of mammograms. Fractal dimension and different methods for the computation of FD are discussed in the next section.

4.2 Fractal Dimension

Mandelbrot in 1982 (Man 1982) has pioneered the use of fractals to describe objects that possess self similarity at all scales and levels of magnification. Fractal objects have irregular shapes and complex structures that cannot be represented adequately by the traditional Euclidean dimension. Fractal Dimension (FD) is the fundamental parameter for depicting fractal characteristics in fractal geometry. It is a number that characterizes the 'structure' of the object. It assigns non integral dimension values to objects that do not suit the traditional Euclidean space of objects.

For example, the dimension of a straight line is unity, but the dimension of a jagged line is a fractional value falling between unity and two, depending on its degree of jaggedness. The fractal dimension has been used in image classification to measure surface roughness where different natural scenes such as mountains, clouds, trees and deserts generate different fractal dimensions. The effective fractal dimension estimation method is a precondition to utilizing fractal dimension to depict fractal characteristic.

The parenchymal and ductal patterns in mammograms possess structures with high local self-similarity which is the basic property of fractal objects (Li 1997). Therefore, fractal method can be applied effectively for the analysis of mammograms.

It is observed that microcalcifications and masses are visible as objects which appear to be added to the mammographic breast background. Some of them are bright, some are faint. Compared with breast background tissue, they have less structure. But the complexity increases in cancerous ones due to the presence of the abnormal tissues. When abnormality in mammograms increases its complexity also increases.

Next section covers the literature survey carried out on the current trends in mammogram image processing and the use of fractals in the analysis of mammograms.

4.3 Literature Survey

There are large numbers of publications available in literature where fractals and fractal based properties are applied for a wide range of applications as seen in chapter 3.

A system for identifying the boundary of liver in Computed Tomography (CT) images was developed by Chen *et. al.* (Cen 1998) using fractal features and deformable contour model. The normalized Fractional Brownian (NFB) motion feature values, the correlation and sum entropy of the spatial gray-level dependence matrices in conjunction with Modified Probability Neural Network (MPNN) were used to discriminate between two types of liver tumors: hepatoma and hemageoma.

Pan and Lin (Pan 2010) classified the normal and cancerous liver tissues using fractal dimension and Probabilistic Neural Network and obtained an accuracy of 92.0% for the test set Wu *et.al* (Wu 1992) developed a new texture feature set called multi-resolution fractal features based on the concepts of multiple resolution imagery and fractional Brownian motion model. The performance of these features was compared with the features like spatial gray-level dependence matrices, the Fourier power spectrum, the gray-level difference statistics, and the Laws' texture energy measures. The new features were able to correctly classify three sets of ultrasonic liver images - normal liver, hepatoma, and cirrhosis and 90% correct classification was observed.

Local fractal dimensions of ECG signal were used by Raghav and Mishra (Rag 2008) as a feature for the classification of ECG arrhythmia.

The fractal features extracted from fractal transformation (FT) by nonlinear interpolation functions and Probabilistic Neural Network (PNN) has been introduced by Lin *et.al* (Lin 2009) to recognize multiple cardiac arrhythmias.

Directional fractal dimension, which measures the degree of roughness along a certain spatial direction along with the Multi-Layer Feed forward Neural Network (MFNN), was used to perform classification of tissue section images of cells from patients suffering from critical limb ischaemia by Shang *et. al.* (Sha 2000). The classifier achieved a classification accuracy of 91%.

The breast ultrasound images were first preprocessed to remove noise by histogram equalization and morphological operations by Chang *et. al.* (Chg 2004) and Chen *et. al.* (Che 2005). The normalized fractional Brownian motion feature vector extracted from the processed images was used for classification with k means classification method. The area under the ROC curve was found to be 0.9218.

A novel method of extraction of Region of Interest (ROI) in the breast ultrasound (BUS) image for the fully automatic classification was shown by Liu *et.al* (Liu 2010). The Area Under the Curve (AUC) of the generated ROIs was obtained as 0.968.

Potlapalli (Pot 1998) developed a new Incremental Fractional Brownian Motion (IFBM) for the classification of textures.

High dimensional biologically inspired feature (BIF) and its variations have been demonstrated by Song and Tao (Son 2010) to be effective and efficient for scene classification.

A new adaptive fuzzy classification algorithm, called influential rule search scheme (IRSS), was developed by Chatterjee and Rakshit (Cht 2004) for automatic construction of the fuzzy membership functions (MFs) and the fuzzy rule base from an input-output data set.

A novel pattern recognition approach was proposed by Baskes *et.al* (Bac2010) based on the Complex Network Theory and complexity analysis. It was illustrated how a shape contour can be effectively represented and characterized as a complex network in a dynamic evolution context, and how degree based measurements can be used to estimate the network complexity through Multi-Scale Fractal Dimension.

For automatic target recognition (ATR) using radar, the local fractal dimensions of a synthetic aperture radar image have been used as features to classify ground targets by Mishra *et. al* (Mis 2007).

The concept of lacunarity and the use of two lacunarity estimation methods (i.e., binary, gray scale) in texture analysis and classification of high resolution urban images were discussed by Myint and Lam (Myi 2005). When compared with the

traditional spectral based classification approach with an accuracy of 55%, lacunarity approaches improved the accuracy dramatically to 92%.

Hadjileontiadis (Had 2009) presented an automatic classification method to discriminate the types of discontinuous breath sounds (DBSs), i.e., fine crackles (FCs), coarse crackles (CC), and squawks (SQ), using lacunarity. They have shown that it provided an efficient discrimination among DBS with a mean classification accuracy of 100%, 99.62%–100%, and 99.75%–100% for the comparison groups of {FC-CC, FC-SQ}, {CC-SQ}, and {FC-CC-SQ}, respectively

A comparative analysis of different feature extraction methods for fingerprint classification based on orientation maps (OMs) and Gabor filters was presented by Rajanna *et.al* (Raj 2009).

The selection of useful features is a fundamental problem in any classification task. Irrelevant and redundant features degrade classification performance. Therefore, Levi and Ullman (Lev 2010) dealt with the goal of selecting a set of features, which is optimum for classification and have minimum redundancy.

Samarabandu *et. al.* (Smb 1993) used the concepts of mathematical morphology to compute the fractal dimension of bone x-rays. This gives an additional advantage of encoding structural information via the selection of a structuring element and also gives a robust texture measure of trabecular bone structures.

Lin *et. al.* (LiK 2001) computed fractal dimension using differential box counting method for extracting eye pairs which achieves an overall hit rate of 100% without head tilt.

It was found by Jiang *et. al.* (Jng 2009) that, the contour fractal dimension as well as the contour and nervure fractal dimension can be used to distinguish between leaves of different types of plants effectively.

The correlation fractal dimension was used by Langi and Kinsner (Lan 1995) as a distinguishing feature for characterizing consonant phonemes that improved the speech recognition performance.

Kinsner and Vera (Kin 2006) classified real world self affine non stationary signals from non linear systems, by the computation of the Variance of the Fractal Dimension Trajectory (VFDT). The features extracted from VFDT were applied to

complex domain neural network and probabilistic neural network gave a classification accuracy of 87%.

A new method for finding the fractal dimension which is less sensitive to sampling frequency was developed by Senevirathne *et.al.* (Sen 1992)

Fractal dimension was used by Fekkai *et. al.* (Fek 2000) to characterize the fluctuations in speech signal and was utilized for the recognition of isolated speech.

The degree of correlation between breast parenchymal patterns was assessed using a simple fractal dimension method by Caldwell *et. al* (Cal 1990). Velanovich (Vel 1998) quantified complex shapes in mammograms by fractal analysis and using this they were classified into benign and malignant with 100% sensitivity and 63% specificity.

The lesions in the mammograms were extracted by SzCkely and Pataki (SzC 2003) by first binarizing the image and then using the shape descriptors derived from moment based and PCA based methods.

Kobatake *et.al* (Kob 1994) enhanced cancerous tumors by the newly developed adaptive iris filter. Shape analysis was then applied to discriminate between malignant tumors and others. For them the average number of false positives per image was only 0.18 while the true positive detection rate was 100%.

Brake and Karssemeijer (Bra 1999) developed three different pixel-based methods for detecting masses based on scale. Their first method utilizes convolution of a mammogram with the Laplacian of a Gaussian, the second method was based on correlation with a model of a mass, and the third was a new one, based on statistical analysis of gradient-orientation maps.

Curvelet transform was applied to classify different types of mammograms by Eltoukhy (Elt 2009). Baeg and Kehtarnavaz (Bae 2002) introduced an automatic CAD system for the classification of mammogram masses into benign and malignant. Here, the two features namely, denseness and architectural distortion were fed to the neural network classifier, which gave an area under the curve of 0.90 in the ROC analysis.

Faye *et. al* (Fay 2009) decomposed mammograms using wavelets to extract a set of coefficients to differentiate between normal and abnormal and then to classify the type of abnormality as benign or malignant tumor.

An appropriate Gabor filter is chosen which can identify the texture differences between normal and abnormal mammograms. To increase classification efficiency and reduce the feature space, statistic t-test and its p- values for feature selection and weighting were proposed by Don and Wang (Don 2009).

Rangayyan *et. al.* (Ran 1997) investigated the use of a new measure of edge strength or acutance of the tumor ROI, to characterize the fuzzy nature of malignant tumor boundaries and the sharply defined nature of benign masses. Acutance alone gave a benign/malignant decision accuracy of about 95% with the MIAS database. They also analyzed the effectiveness of shape factors like compactness, in distinguishing between circumscribed and spiculated tumors, achieving an accuracy of 92.3%.

Khuwaja and Reza (Khu 2004) proposed a bi-modal artificial neural network (ANN) for breast cancer classification system. The microcalcifications are extracted with adaptive learning vector quantization networks that are trained with cancer/malignant and normal/benign breast digital mammograms. The performance of the networks is evaluated using ROC curve analysis which gave a sensitivity-specificity of 98.0-100.0 for the CC view and 96.0-100.0 for the MLO view.

Since fractals possess properties, it can be easily approximated to physiological entities. Therefore to have a better classification accuracy fractal dimension and fractal based features are used in this research.

The three fractal dimension estimation methods present in literature are discussed in the next section.

4.4 Fractal Dimension Estimation Methods

In this research, three methods for computing fractal dimension were considered. They are the Differential Box Counting, Blanket and Triangular Prism Surface Area methods. The discussion starts with the conventional box counting method.

4.4.1. Box Counting Method

The simplest method to compute fractal dimension is the box counting method, which is based on the concept of self-similarity. In a Euclidean n -space, a bounded set A , is said to be self-similar when A is the union of N_r distinct (non overlapping) copies of itself, each of which has been scaled down by a ratio of r . The fractal dimension D is related to the number N_r and the ratio r as follows:

$$D = \lim_{r \rightarrow 0} \frac{\log(N_r)}{\log(1/r)} \quad (4.1)$$

where N_r is the minimum number of distinct fractal copies of A in the scale r i.e. the number of boxes of size r . The union of N_r of all the distinct copies must cover the entire set A completely.

The main difficulty with this method is that real world images are seldom self similar. Also, this method is appropriate for finding the fractal dimension of binary images only. Hence, the box counting method was modified and made suitable to be applied to gray level images.

4.4.2 Differential Box Counting Method

Sarkar and Chaudhuri (Sar 1994) proposed the differential box counting (DBC) method by modifying the box counting method and have compared it with other conventional methods. Consider an image of size $M \times M$ pixels. Assume that the image is represented in a 3D space, with (x, y) axes denoting the spatial coordinates, while the z axis denoting the gray level. The fig.4.3 shows the image plane and the image intensity surface for estimating the fractal dimension using the differential box counting method.

The (x, y) space is partitioned into grids of size $s \times s$, where $M/2 \geq s > 1$, is an integer. Then, $r = s/M$. On each grid there is a column of boxes of size $s \times s \times s'$, where s' is the side along the z direction corresponding to the gray level axis. If the total number of gray levels is G , then, $\lfloor G/s' \rfloor = \lfloor M/s \rfloor$. Numbers from 1, 2, ... are

assigned to the boxes starting from the lowest gray level value. Let the minimum and the maximum gray level of the image in the $(i, j)^{th}$ grid fall in box number k and l , respectively.

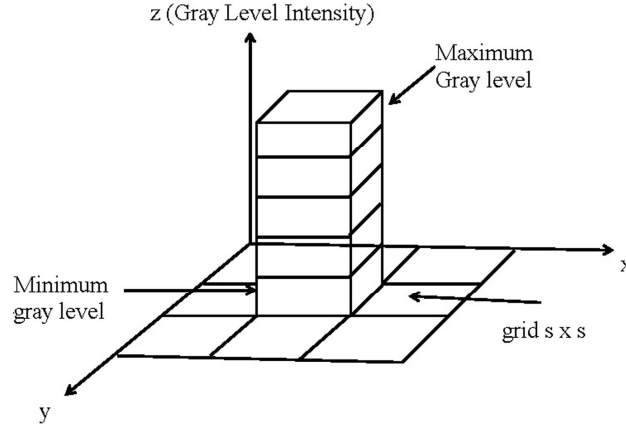


Fig 4.3 Schematic for finding the FD using Differential Box counting method

The contribution of N_r in $(i, j)^{th}$ grid is given by:

$$n_r(i, j) = l - k + 1 \quad (4.2)$$

Due to the differential nature in computing n_r this method is called differential box counting method. The contributions from all grids are found by:

$$N_r = \sum_{i,j} n_r(i, j) \quad (4.3)$$

N_r is computed for different values of s i.e. different values of r . Using equation (4.2) D , the fractal dimension can be estimated, from the least square linear fit of $\log(N_r)$ against $\log(1/r)$. A random placement of boxes is applied in order to reduce quantization effects.

4.4.3 Blanket Method

Peleg *et al* (Pel 1984) used the blanket method approach of measuring the fractal dimension. Image can be viewed as a hilly terrain surface whose height from the normal ground is proportional to the image gray value as shown in fig. 4.4.

A small portion of mammogram is taken out. The z axis gives the gray level. Gray levels of this region vary from 185 to 230. The upper gray levels are denoted by red color while the lower gray levels are given blue color. Then all points in the three dimensional space at distance ε from the surface on both sides create a blanket of thickness 2ε .

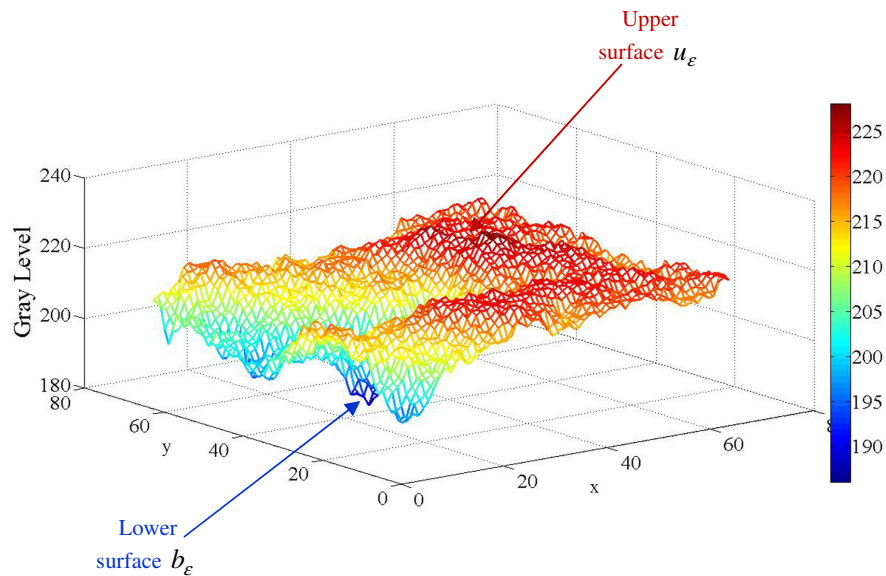


Fig 4.4 Gray Level of Mammogram

The estimated surface area is the volume occupied by the blanket divided by 2ε . All points in the three dimensional space at distance ε from the surface are considered, covering the surface with a "blanket" of thickness 2ε . The covering blanket is defined by its upper surface u_ε and lower surface b_ε .

Initially, i.e. when $\varepsilon = 0$, the upper and lower surfaces are given by the same gray level function $g(i, j)$, i.e.

$$u_0(i, j) = b_0(i, j) = g(i, j) \quad (4.4)$$

For $\varepsilon = 1, 2, \dots$ the blanket surfaces are defined as:

$$u_\varepsilon(i, j) = \max \left\{ u_{\varepsilon-1}(i, j) + 1, \max_{|(m,n)-(i,j)| \leq 1} u_{\varepsilon-1}(m, n) \right\} \quad (4.5)$$

$$b_\varepsilon(i, j) = \min \left\{ b_{\varepsilon-1}(i, j) - 1, \min_{|(m,n)-(i,j)| \leq 1} b_{\varepsilon-1}(m, n) \right\} \quad (4.6)$$

A point (x, y, f) will be included in the blanket of ε when $b_\varepsilon(x, y) < f \leq u_\varepsilon(x, y)$.

The blanket definition uses the fact that the blanket of the surface for radius ε includes all the points of the blanket for radius $\varepsilon - 1$, together with all the points within radius 1 from the surfaces of that blanket. Expressions (4.5) and (4.6) ensure that the new upper surface and lower surfaces are higher/lower by at least 1 from $u_{\varepsilon-1}/b_{\varepsilon-1}$, and also at distance at least 1 from $u_{\varepsilon-1}/b_{\varepsilon-1}$ in the horizontal and vertical directions.

The volume of the blanket is computed from u_ε and b_ε by the following equation:

$$V_\varepsilon = \sum_{i,j} (u_\varepsilon(i, j) - b_\varepsilon(i, j)) \quad (4.7)$$

The surface area of the blanket can be measured from the volume as:

$$A_\varepsilon = \frac{V_\varepsilon - V_{\varepsilon-1}}{2} \quad (4.8)$$

Another definition for the surface area is $V_\varepsilon / 2\varepsilon$. This is necessary, since V_ε depends on all small scales features. Subtracting $V_{\varepsilon-1}$ isolates just those features that change from scale $\varepsilon - 1$ to scale ε . When a pure fractal object is analyzed, both definitions are identical since property changes are independent on scale, and measurements between any two different scales will yield the same fractal dimension. However, for non fractal objects, this isolation from the effects of smaller scale features is necessary. The definition in the above equation (4.8) gives reasonable measures for both fractal and non fractal surfaces.

According to Mandelbrot (Man 1982), the area of the fractal surface is:

$$A_\varepsilon \approx F\varepsilon^{2-D}, \varepsilon = 1, 2, \dots \quad (4.9)$$

where D is the fractal dimension. When plotting A_ε versus ε on a log-log scale, a straight line is obtained with slope $2 - FD$. This curve is not a straight line for non fractal surfaces. The slope of the best fitting straight line gives the fractal signature $S(\varepsilon)$ (discussed in detail in section 4.4.3.1). For fractal objects $S(\varepsilon)$ should be equal to $2-D$ for all ε .

$$\text{i.e. Fractal Signature } S(\varepsilon) \text{ or Slope} = 2 - FD \quad (4.10)$$

According to Tao *et. al* (Tao 2000) the fractal dimension can be computed from equation (4.9) as follows. A line can be drawn when any two points are known. Therefore, two values ε and $\varepsilon - 1$ are used for the computation of fractal dimension as:

$$A_{\varepsilon_1} \approx F\varepsilon_1^{2-D}, \quad A_{\varepsilon_2} \approx F\varepsilon_2^{2-D}$$

or,

$$\frac{A_{\varepsilon_1}}{A_{\varepsilon_2}} \approx \frac{\varepsilon_1^{2-D}}{\varepsilon_2^{2-D}}$$

Taking logarithm on both sides:

$$2 - D \approx \frac{\log_2(A_{\varepsilon_1}) - \log_2(A_{\varepsilon_2})}{\log_2(\varepsilon_1) - \log_2(\varepsilon_2)}$$

$$D \approx 2 - \frac{\log_2(A_{\varepsilon_1}) - \log_2(A_{\varepsilon_2})}{\log_2(\varepsilon_1) - \log_2(\varepsilon_2)} \quad (4.11)$$

4.4.3.1 Fractal Signature

The magnitude of the fractal signature $S(\varepsilon)$, is related to the amount of detail included on the blanket of size ε . High value of the signature $S(\varepsilon)$, is associated with large gray level variations at distance ε . High value of $S(\varepsilon)$ at small ε is due to high frequency gray level variations while high values for larger ε result from major low frequency variations. Thus the fractal signature $S(\varepsilon)$ gives important information about the fineness of the variation of the gray level surface.

The different types of mammograms are compared based on differences between their fractal signatures. The distance for two different images with signatures $S_i(\epsilon)$ and $S_j(\epsilon)$ is defined by:

$$D(i, j) = \sum_{\epsilon} (S_i(\epsilon) - S_j(\epsilon))^2 \log \left(\frac{\epsilon + \frac{1}{2}}{\epsilon - \frac{1}{2}} \right) \quad (4.12)$$

The weighting factor of $\log[(\epsilon + 0.5)/(\epsilon - 0.5)]$ is due to the unequal spacing of the points in the log-log scale.

4.4.3.2 Differential Fractal Signatures and Distance Measurement

Consider an image of light particles scattered over a dark background. Since the high gray level value stands for white, the expression for minimum in equation (4.6) will shrink the light regions, and the rate of this shrinking will only depend on the shape property of the particles. The maximum operator in equation (4.5) however will shrink the background regions and the rate of this shrinking will mainly be affected by the distribution of particles. To consider this asymmetry, the surface area measurement is divided into two parts: measuring the area of the gray level surface when viewing from “above” and measuring the area when viewing the surface from “below”. Thus, the definition of volume in equation (4.7) can be changed to the following two definitions of “upper volume” V^+ and “lower volume” V^- as follows:

$$V_{\epsilon}^+ = \sum_{i,j} u_{\epsilon}(i, j) - g(i, j) \quad (4.13a)$$

and

$$V_{\epsilon}^- = \sum_{i,j} g(i, j) - b_{\epsilon}(i, j) \quad (4.13b)$$

Similarly the expression for area in (4.8) is also changed into “top area” A^+ and “bottom area” A^- as follows:

$$A_{\epsilon}^+ = \frac{V_{\epsilon}^+ - V_{\epsilon-1}^-}{2} \quad (4.14a)$$

and

$$A_{\epsilon}^- = \frac{V_{\epsilon}^- - V_{\epsilon-1}^-}{2} \quad (4.14b)$$

The two different fractal signatures S^+ and S^- are computed as before. The S^- plot represents the shape of the parts while the S^+ graphs represent the background of the image.

Analyzing again, now by differentiating between the top and bottom areas, the differential distance D' between two textures i and j are defined as:

$$D'(i, j) = \sum_{\varepsilon} \left\{ \left[(S_i^+(\varepsilon) - S_j^+(\varepsilon))^2 + (S_i^-(\varepsilon) - S_j^-(\varepsilon))^2 \right] \cdot \log \left(\frac{\varepsilon + \frac{1}{2}}{\varepsilon - \frac{1}{2}} \right) \right\} \quad (4.15)$$

4.4.4 Triangular Prism Surface Area Method

The third method to measure fractal dimension used in this research work is the Triangular Prism Surface Area (TPSA) method proposed by Clarke (Cla 1986). The schematic representation for the measurement of the triangular prism surface area is shown in fig. 4.5.

The original image is assumed to be of size $M \times M$ as in the above method. The steps required for finding fractal dimension using TPSA method, discussed by Tang and Wang (TaM 2005) are given below:

Step 1: The image is divided into different grids of size r . Let the four points of every square grid correspond to four points A, B, C, D on the fractal surface. These points are represented by the gray level value at that point. i.e. height of this grid corresponds to the gray level values $h(i, j)$, $h(i, j+1)$, $h(i+1, j)$ and $h(i+1, j+1)$ respectively.

Step 2: The distance from the ground to the center of each grid cell of the four heights of the adjacent points can be calculated as:

$$h_0 = \frac{1}{4} [h(i, j) + h(i, j+1) + h(i+1, j) + h(i+1, j+1)] \quad (4.16)$$

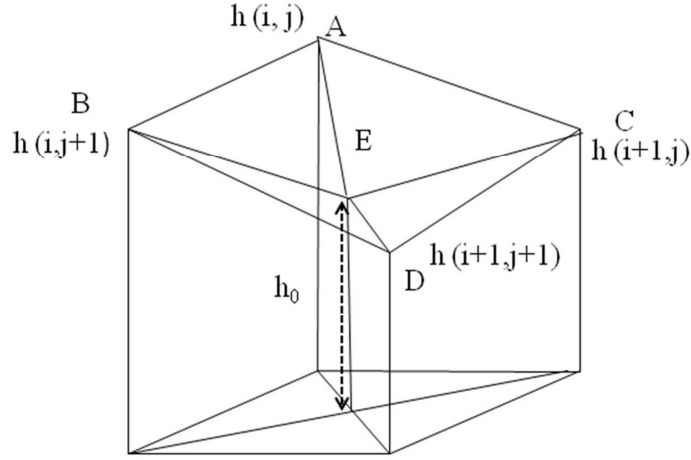


Fig. 4.5 Schematic for finding TPSA

Step 3: The top of the grid is divided into four triangles ABE , ACE , CDE and BDE .

The area of the triangle ABE is determined as:

$$S_{ABE} = \sqrt{l_1(l_1 - a_1)(l_1 - b_1)(l_1 - c_1)} \quad (4.17)$$

$$\text{where } l_1 = \frac{1}{2}(a_1 + b_1 + c_1) \quad (4.18)$$

$$a_1 = \sqrt{[h(i, j) - h(i, j+1)]^2 + r^2} \quad (4.19a)$$

$$b_1 = \sqrt{[h(i, j) - h_0]^2 + 0.5r^2} \quad (4.19b)$$

$$c_1 = \sqrt{[h(i, j+1) - h_0]^2 + 0.5r^2} \quad (4.19c)$$

Step 4: The area of the remaining triangles ACE , CDE and BDE are also found out similarly. Thus, the approximate real area of a fractal surface in a given grid cell with a scale of $r \times r$ is given by:

$$S(i, j) = S_{ABE} + S_{ACE} + S_{CDE} + S_{BDE} \quad (4.20)$$

Step 5: Considering the entire image, the total area of the fractal surface is:

$$S(r) = \sum_{i,j=1}^{N(r)} S(i, j) \quad (4.21)$$

where $N(r)$ is the total number of the regular squares of size $r \times r$.

Step 6: In fractal geometry, the total area of the fractal surface $S(i, j)$, the scale δ and the fractal dimension D are related by

$$s(r) \sim r^{2-D} \quad (4.22)$$

Repeat steps 1-5 with different values of r . Then $\log(S(r))$ and $\log(r)$ is plotted in the log-log co-ordinate system. If the slope of the best fitting straight line joining these points is b , the fractal dimension D of the image is:

$$D = 2 - b \quad (4.23)$$

4.5 Fractal Features

The main problem with the fractal dimension approach is that it cannot uniquely characterize the texture pattern. Different textures may have the same fractal dimension. This is due to the combined differences in coarseness and directionality i.e. dominant orientation and degree of anisotropy (Man 1982). Hence features based on fractal dimension were considered. Five features based on fractal dimension used in texture segmentation (Chd 1995) were tried for characterizing mammograms. They are the FD of the original image, the high gray valued image, the low gray valued image, the horizontally smoothed image and vertically smoothed image. In addition to these features a new fractal feature was derived from the average of four pixels of the image. The different fractal features utilized in this research are discussed below.

4.5.1 Fractal Feature 1 (f_1)

The FD of the original image I_1 is computed using overlapping windows of size $(2W + 1) \times (2W + 1)$. Thus, at point (i, j) the first feature value $F_1(i, j)$ is defined as

$$F_1(i, j) = FD\{I_1(i+l, j+k); -W \leq l, k \leq W\} \quad (4.24)$$

where, FD is the fractal dimension computed using any of the methods described in section 4.4. Since the fractal dimension is greater than the topological dimension, the value of F_1 is between 2 and 3. The normalized feature is defined as $f_1 = F_1(i, j) - 2$, such that $0 \leq f_1 \leq 1$. Thus all the normalized fractal features are between 0 and 1.

4.5.2. Fractal Features 2 and 3 (f_2 and f_3)

The two modified images called high and low gray-valued images, I_2 and I_3 respectively are defined as:

$$I_2(i, j) = \begin{cases} I_1(i, j) - L_1, & \text{if } I_1(i, j) > L_1 \\ 0 & \text{otherwise} \end{cases} \quad (4.25)$$

$$I_3(i, j) = \begin{cases} 255 - L_2, & \text{if } I_1(i, j) > (255 - L_2) \\ I_1(i, j) & \text{otherwise} \end{cases} \quad (4.26)$$

where

$$L_1 = g_{\min} + av/2; \quad (4.27a)$$

$$L_2 = g_{\max} - av/2; \quad (4.27b)$$

with g_{\max} , g_{\min} and av denoting the maximum, minimum and average gray value in the original image I_1 , respectively. Even if two images I_1 and J_1 have the same fractal dimension, their high gray-valued images I_2 and J_2 may not have an identical roughness and so their FDs would be different. The same holds for I_3 and J_3 . The normalized fractal features f_2 and f_3 are computed from I_2 and I_3 similar to the computation of f_1 from I_1 .

4.5.3 Fractal Feature 4 and 5 (f_4 and f_5)

Fractal dimension of an image is directly related to its roughness and therefore its value will be reduced by gray level smoothing. If the texture is smoothed along the direction of its dominant orientation, the FD will be affected least for a highly oriented

texture. But when the smoothing direction is perpendicular, the FD will be considerably reduced. A texture having a low degree of anisotropy will show an identical effect on the FD, irrespective of the smoothing direction.

Images can be smoothed in the horizontal and vertical direction as:

$$I_4(i, j) = \frac{1}{2W + 1} \sum_{k=-W}^W I(i, j + k) \quad (4.28)$$

$$I_5(i, j) = \frac{1}{2W + 1} \sum_{k=-W}^W I(i + k, j) \quad (4.29)$$

W is the same as in fractal feature f_1 . The normalized FD features f_4 and f_5 are computed similar to that of f_1 .

4.5.4 Fractal Feature 6 (f_6)

A new fractal feature is derived from the smoothed image obtained by computing the average of four neighboring pixels using a non overlapping 4×4 window. The new image is given by:

$$I_6(i, j) = \frac{1}{4} \sum_{i=1}^2 \sum_{j=1}^2 I(i, j) \quad (4.30)$$

Here, four pixels in the image were replaced with a single pixel. Thus the size of the image will be reduced by 4 (i.e. $M \times M$ image is reduced to $M/2 \times M/2$). This reduces the difference between the minimum and maximum gray levels in the new image. This in turn, reduces the variance of the gray levels in the prism (in fig. 4.5). So when the area of the triangles formed from the prism is calculated, in the TPSA method, which is used for evaluating f_6 , the area will be reduced. This means that the spread of the area, which is used to calculate, f_6 will be reduced while considering regions of mammogram containing normal and cancerous tissues. Thus, when the feature f_6 is calculated, the overlap between the values for the different categories can be avoided. This overlapping of feature values among the different categories was the problem with all the fractal features from f_1 to f_5 . Feature f_6 is novel and features f_1 to f_5 are used for the comparison of classification result. A good feature should remain unchanged with changes within the class and should reveal important

differences when discriminating between patterns of different classes (Bas 2004) and this is satisfied by the results obtained with f_6 .

4.6 Conventional features used for the comparison with fractal features

The performance of the fractal features in discriminating the different classes of mammograms was compared with the following conventional features based on intensity, texture, etc.

4.6.1 Statistical Descriptors

- a. **Mean:** The mean of a random variable is defined using the probability density or mass function. It provides a measure of central tendency of the distribution. (Mtz 2008)

$$\mu = \sum_{k=1}^N \frac{f_k n_k}{n} \quad (4.31)$$

where N denotes the gray levels of the mammogram, f_k is the k^{th} gray level and n_k is the number of pixels having gray level f_k and n is the total number of pixels in the region considered (Min 2003a)

- b. **Variance:** The variance of a random variable is a measure of the dispersion in the distribution. If the random variable has a large variance, then the observed value of the random variable is more likely to be far from the mean μ . The square root of variance is called standard deviation (Mtz 2008).

$$\sigma^2 = \sum_{k=1}^N \frac{(f_k - \mu)^2 n_k}{n} \quad (4.32)$$

- c. **Skewness:** It is a measure of the asymmetry of the data around the sample. If the skewness is negative, the data is spread out more to the left of the mean than to the right. If skewness is positive, the data are spread out more to the right. The skewness of the normal distribution or any perfectly symmetric distribution is zero (Min 2003a).

$$\mu_3 = \frac{1}{\sigma^2} \sum_{k=1}^N \frac{(f_k - \mu)^3 n_k}{n} \quad (4.33)$$

- d. **Kurtosis:** Kurtosis is defined as the normalized form of the fourth central moment of a distribution and it indicates the degree of peakedness of a distribution. It is based on the size of the tail of the distribution. It is a classical measure of non gaussianity and can be positive or negative. Distributions with relatively large tails have a negative kurtosis and are called subgaussian or leptokurtic. Those with small tails and positive kurtosis are called super Gaussian or platykurtic. A distribution with the same kurtosis as the normal distribution is called mesokurtic. The kurtosis of a normal distribution is zero.

$$\mu_4 = \frac{1}{4} \sum_{k=1}^N \frac{(f_k - \mu)^4 n_k}{n - 3} \quad (4.34)$$

Kurtosis is widely used as a measure of non gaussianity because of its computational and theoretical simplicity. However, its value can be very sensitive to outliers. Its value may depend on only few observations in the tails of the distribution, which may be erroneous or irrelevant (Min 2003a).

4.6.2 Textural features

Texture features are used in a large number of applications like image analysis, segmentation and classification. These features are particularly valuable in medical image processing. (Mir 1995, Wu 1992). Two basic methods for texture description exist: statistical and structural. Statistical methods employ features extracted from the image which measure coarseness, contrast, directionality and other textural characteristics. Structural methods describe texture by means of primitive descriptions and primitive placement rules (Min 2003a).

The Spatial Gray Level Dependence Method (SGLDM) is concerned with the spatial distribution and spatial dependence among the grey levels in a local area (Mir 1995). The SGLD method is based on the estimation of the second order joint conditional probability density functions,

$$f(i, j|d, \theta), \text{ where } \theta = 0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}, 180^{\circ}, 225^{\circ}, 270^{\circ}, 315^{\circ} \quad (4.35)$$

Each $f(i, j|d, \theta)$ is the probability of going from gray level i to gray level j . It is constructed by counting the number of occurrences of pixel pairs at a given displacement d for a given direction θ . The co occurrence matrix is computed by specifying a distance between the pixels in each pair (displacement vector) and whether the direction of the vector is important or not. In particular, to compute the SGLD matrix for an image $I(i, j)$, a displacement vector $d=(x, y)$ is defined. The $(i, j)^{\text{th}}$ element of the SGLD matrix:

$$s(i, j|d, \theta) = \frac{N(i, j|d, \theta)}{\sum_{i, j} N(i, j|d, \theta)} \quad (4.36)$$

$N(i, j|d, \theta)$ is the number of occurrences of gray levels i and j at a distance d in $I(i, j)$. The matrix is then normalized so that it can be treated as probability density function.

If a texture is coarse and d is small compared to the sizes of the texture elements, the pairs of points at the inter sample distance should usually have similar gray levels. This means that the probability distribution in the matrix is concentrated on or near its diagonal. On the other hand, for a fine texture, the gray levels of the points separated by the distance should be quite different so that the probability distribution is distributed away from its diagonal.

The most significant disadvantage of the co-occurrence matrix is its dependency on the number of gray levels in the entire range ((Mir 1995, Wu 1992)). Since texture is usually measured in a small region, a large number of entries are zero contributing nothing to the texture description of the region. The computational time for the texture feature extraction operations include the time for processing these also. The above problems become more serious when the examined images are composed of a large number of gray levels (Min 2003a, The 2006).

In this work, the following features are evaluated from the SGLD matrix for a fixed d and θ .

- a. **Contrast:** The contrast feature is a measure of the contrast or the amount of local variations present in an image and is given by:

$$\sum_n n^2 \sum_{i,j:n=|i-j|} s(i, j) \quad (4.37)$$

b. **Correlation**

$$\sum_{i,j} \frac{s(i, j)ij - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (4.38)$$

Correlation shows the degree of dependency between the pixels.

c. **Cluster Shade**

$$\sum_i \sum_j (i + j - \mu_i - \mu_j)^3 s(i, j) \quad (4.39)$$

d. **Cluster Prominence**

$$\sum_i \sum_j (i + j - \mu_i - \mu_j)^4 s(i, j) \quad (4.40)$$

e. **Dissimilarity**

Euclidean distance was used to measure similarity

f. **Energy**

$$\sum_{i,j} s(i, j)^2 \quad (4.41)$$

g. **Entropy**

$$- \sum_{i,j} s(i, j) \log s(i, j) \quad (4.42)$$

It is a measure of randomness. For smooth images it takes low values (Bas 2004)

h. **Homogeneity**

$$\sum_{i,j} \frac{1}{1 + (i - j)^2} s(i, j) \quad (4.43)$$

i. **Maximum probability**

$$\max_{ij} P_{ij} \quad (4.44)$$

j. **Sum average**

$$\sum_{i=2}^{2N_g} i s_{x+y}(i) \quad (4.45)$$

k. **Sum variance**

$$\sum_{i=2}^{2N_g} i + \sum_{j=2}^{2N_g} s_{x+y}(j) \log(s_{x+y}(j)) \quad (4.46)$$

l. **Sum entropy**

$$\sum_{i=2}^{2N_g} s_{x+y}(i) \log(s_{x+y}(i)) \quad (4.47)$$

m. **Difference variance**

$$\text{Variance of } s_{x-y} \quad (4.48)$$

n. **Difference entropy**

$$- \sum_{i=1}^{N_g} s_{x+y}(i) \log\{s_{x-y}(i)\} \quad (4.49)$$

o. **Information measure of correlation**

$$\frac{HXY - HXY1}{\max\{HX, HY\}} \quad (4.50)$$

p. **Inverse difference moment normalized**

$$\sum_i \sum_j \frac{1}{1 + (i - j)^2} s(i, j) \quad (4.51)$$

Describes the local homogeneity of an image (Bas 2004).

4.7 Statistical Analysis

The performances of these features in classifying the mammogram were evaluated based on the Receiver Operating Characteristics (ROC) (Faw 2006, Han 1982, Obu 2005). An ROC graph is a technique for visualizing, organizing and selecting classifiers based on their performance (Faw 2006). At the same time, a method to specify the performance of a classifier is in terms of its ROC curve. (Bas 2004). An ROC (Obu 2005) curve is a plot of a test's false positive rate (FPR) or (1-specificity), plotted on the horizontal axis, versus its sensitivity (plotted on the vertical axis). The plotted co-ordinates are connected with line segments to construct an empiric ROC. Each point on the curve represents the sensitivity and FPR at a different threshold.

An ROC curve begins at the (0, 0) co-ordinate, corresponding to the strictest decision threshold whereby all test results are negative for disease. The ROC ends at the (1,1) co-ordinates corresponding to the most lenient decision threshold whereby all test results are positive for disease.

The line connecting the (0, 0) and (1, 1) co-ordinates is called the “chance diagonal” and represents the ROC curve of diagnostic test with no ability to distinguish patients with and without disease. The farther away an ROC curve from the chance diagonal, the closer to the upper left hand corner, the better the discriminating power and the diagnostic accuracy of the test is.

The area under the curve (AUC) is an accepted modality for comparing classifier performance. A perfect classifier has TP rate of 1.0, and an FP rate of 0.0, resulting in an AUC of 1.0. The most popular measure of accuracy is the Area under the ROC curves denoted by AUC which ranges from the value 0.5 (chance) to 1.0 (perfect discrimination or accuracy). Another measure to summarize accuracy is the Youden’s index, defined as sensitivity + specificity-1. There is another measure called the accuracy. It is shown that the probability of a correct diagnosis is equivalent to

Probability (correct diagnosis) = $PREV_s \times \text{sensitivity} + (1 - PREV_s) \times \text{specificity}$. where $PREV_s$ is the prevalence of disease in the sample.

Thus an ROC demonstrates several things:

- Shows the trade-off between sensitivity and specificity.
- The closer the curve approaches to the left hand border of the top border of the ROC space, the more accurate the test.
- The closer the curve comes to the 45⁰ diagonal of the ROC space, the less accurate the test.
- The AUC for an ideal classifier is 1 (Bas 2004).

4.8 Implementation of Classification of Mammograms using various fractal features

The various features mentioned above were evaluated and the results are mentioned in the following sections.

4.8.1 Database used

The mammograms used for validating the above discussed methods were obtained from the freely available two standard online mammogram research databases namely, Mammogram Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM).

4.8.1.1 MIAS Database

The images in the MIAS database are digitized at a resolution of 50 μm per pixel, with 1,024 \times 1,024 pixel size and at 256 gray levels (Suc 1994). The accompanied “ground truth” contains details regarding the character of the background tissue, class, and severity of the abnormality and x, y coordinate of its center and radii. The size of microcalcification generally varies from 50 to 1,000 μm . The smallest microcalcification is equivalent to 1 pixel of the image. The subtlety rating of these mammograms are found to be 1, 2, and 3 as per (Pis 1998) which indicates that the lesions are detectable only by an expert mammographer, likely to be detected by an expert, and likely to be detected by observer with good mammographic training, respectively.

4.8.1.2 DDSM Database

The DDSM contains mammograms obtained from Massachusetts General Hospital, Wake Forest University School of Medicine, Sacred Heart Hospital and Washington University of St. Louis School of Medicine. The four standard views (medio-lateral oblique and cranio caudal) from each case were digitized on one of four different.

The cases were all from mammography examinations conducted between October of 1988 and February of 1999.

The cases were assigned to volumes according to the severity of the finding. ‘Normal’ volumes contain mammograms from screening exams that were read as ‘normal’ and had a normal screening exam four years later (plus or minus 6 months). ‘Benign without callback volumes’ contain examined mammograms that had an abnormality that was noteworthy but did not require the patient to be recalled for any additional workup. ‘Benign’ volumes contain cases in which something suspicious was found and the patient was recalled for some additional work-up that resulted in a benign finding. ‘Cancer’ volumes contain cases in which a histologically proven cancer was found. Each volume may contain cases that include less severe findings in addition to the more severe findings that resulted in the assignment of a case to a particular volume.

Every case in the DDSM contains the patient age, the screening exam date, the date on which the mammograms were digitized and the ACR breast density, that was specified by an expert radiologist. Cases in all volumes other than the normal volume contain pixel level ground truth markings of abnormalities. Each marking contains a subtlety value and a description that was specified by an expert mammography radiologist using the BI-RADS™ (ACR 1998) lexicon (Hea 2001).

The following table 4.1 shows the number of mammograms from the two databases used for the study.

Table 4.1 No of Mammograms of each class obtained from the MIAS and DDSM Database used for the study

Type of Mammogram	MIAS	DDSM
Normal	166	180
Masses	Benign	54
	Malignant	39
Microcalcifications	Benign	13
	Malignant	15
Total	287	512

4.9 Results and Discussions

Different Regions of Interest (ROI) viz. 64×64 , 128×128 , 256×256 , were chosen from the mammogram, depending on the radius of the cancerous region present in the image. In the normal mammograms also same size for regions of interest were considered.

4.9.1 Evaluation using fractal signatures and distance measures

The fractal signatures computed according to sections 4.4.3.1 and 4.4.3.2 for the different class of mammograms are shown in the table 4.2. The values of the signatures S , S^+ and S^- obtained for normal mammograms, mammograms with microcalcifications and masses which are both benign and malignant are shown in the table. For normal mammograms the average values of the signatures are the lowest, with 0.0887, 0.0863 and 0.0632 for S , S^+ and S^- respectively. But, the range of values for one class overlap with the values of other classes, for all the fractal signatures S , S^+ and S^- and hence cannot be used efficiently for classification of mammograms. The average standard deviation and range of these values is also given in the table 4.2.

Next, the distance D between the different classes of mammograms based on equation (4.12) was measured. The table 4.3 shows the average value of the distance D obtained between the different classes of mammograms. It is seen from the table that there is sufficient distance between different classes, except for normal and malignant microcalcifications, which is found to be very less of 0.0556. The distance between normal and other classes should be high because when abnormality is present, it should be detected correctly. Hence, this distance measure D cannot be used as a good classifier for characterizing mammograms.

Table 4.2
Comparison of different fractal signatures obtained for the different classes of mammograms

Mammograms	Fractal Signatures								
	Signature S			Signature S+			Signature S-		
	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev
Normal	0.0347 -	0.0887	0.0051	0.0323 -	0.0863	0.0197	0.0291 -	0.0632	0.0015
Mass	0.1267 -			0.0967 -			0.0863 -		
Benign	0.1165 -	0.1347	0.134	0.197- 0.382	0.2915	0.026	0.1573 -	0.3743	0.1068
Malignant	0.3439 -			0.168- 0.356	0.274	0.0839	0.4382 -		
Micro Calcifications	0.2709 -	0.3135	0.135	0.203- 0.342	0.2258	0.021	0.149- 0.385	0.2733	0.0547
Benign	0.3722 -			0.0684 -	0.2569	0.112	0.155- 0.286	0.1988	0.016
Malignant	0.3906 -	0.7087	0.0285	0.658- 0.8319	0.3445	0.0095	0.121- 0.661	0.3553	0.3118

Table 4.3
Average value of Distance D between the different classes of mammograms

	Normal	Mass (Benign)	Mass (Malignant)	Calcifications (Benign)	Calcifications (Malignant)
Normal	0	0.2793	0.1137	0.55	0.0556
Mass (Benign)	0.2793	0	0.154	0.169	0.0713
Mass (Malignant)	0.1137	0.154	0	0.158	0.1033
Calcifications (Benign)	0.55	0.169	0.158	0	0.178
Calcifications (Malignant)	0.0556	0.0713	0.1033	0.178	0

Table 4.4 illustrates the differential distance D' computed using equation (4.15). There is ample variation in the differential distance measure between the normal and the abnormal mammograms.

Table 4.4
Differential Distance D' between the different classes of mammograms

	Normal	Mass (Benign)	Mass (Malignant)	Calcifications (Benign)	Calcifications (Malignant)
Normal	0	0.583	0.3171	0.213	0.2189
Mass (Benign)	0.583	0	0.2833	0.1311	0.6198
Mass (Malignant)	0.3171	0.2833	0	0.588	0.2702
Calcifications (Benign)	0.213	0.1311	0.588	0	0.552
Calcifications (Malignant)	0.2189	0.6198	0.2702	0.552	0

But the differential distance D' cannot identify exactly to which class a mammogram belongs. Therefore, it was not possible to generalize and categorize the class of mammograms using this distance measure D' .

4.9.2. Evaluation of Fractal Dimension Estimated from different methods

The fractal dimensions of the mammograms were calculated using the Triangular Prism Surface Area method (TPSA), Differential Box Counting (DBC) method and the blanket method. For an ROI of $M \times M$, in the TPSA method and the DBC methods an overlapping grid size of 1 to M were considered. For the blanket method, blanket size ϵ was varied from 0 to 20. The results obtained are given in table 4.5.

The presence of breast cancer increases the irregularity in the breast tissue, which increases its fractal dimension. As normal mammograms have a homogeneous structure when compared to the diseased ones, its fractal dimension (FD) should be less. In all the three FD computation methods, as shown in table 4.5, the FD of the normal mammograms was found to be the least. The range of values of fractal dimension overlaps each other for the various classes in the case of the DBC and Blanket methods. For the TPSA method there is overlap between the FD values for benign and malignant microcalcifications and a few malignant microcalcifications.

Table 4.5
Comparison of the fractal dimensions obtained by TPSA, DBC and Blanket methods

Fractal Dimension computed using										
Mammograms	TPSA			DBC			Blanket			
	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	
Normal	2.105	0.0469	2.027-2.176	2.132	0.0714	2.01-2.362	2.1864	0.1036	2.0201-2.3641	
Masses	Benign	2.313	0.079	2.179-2.476	2.226	0.09	2.067-2.397	2.232	0.1141	2.075-2.487
	Malignant	2.848	0.0765	2.707-2.986	2.346	0.163	2.129-2.635	2.4558	0.1634	2.1426-2.6995
Micro calcifns.	Benign	2.593	0.0585	2.482-2.688	2.140	0.0334	2.08-2.193	2.416	0.0511	2.3563-2.5798
	Malignant	2.6141	0.0783	2.502-2.764	2.182	0.0240	2.116-2.215	2.35	0.1631	2.1125-2.5777
Classification Accuracy %		80.17			6.436			6.436		

20 Sample values of the fractal dimension obtained by the TPSA for the different types of mammograms are given in the table 4.6

Table 4.6
Sample Fractal Dimension obtained for different mammograms using TPSA method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	2.06	2.476	2.841	2.606	2.644
2	2.116	2.299	2.713	2.557	2.538
3	2.027	2.179	2.947	2.688	2.676
4	2.094	2.451	2.942	2.647	2.523
5	2.086	2.402	2.804	2.6027	2.615
6	2.105	2.394	2.754	2.564	2.592
7	2.152	2.357	2.824	2.546	2.502
8	2.085	2.289	2.952	2.579	2.616
9	2.04	2.208	2.942	2.65	2.568
10	2.175	2.37	2.788	2.482	2.683
11	2.145	2.225	2.857	2.565	2.604
12	2.075	2.303	2.986	2.669	2.585
13	2.052	2.298	2.897	2.547	2.6075
14	2.169	2.385	2.857	2.55	2.567
15	2.052	2.315	2.751	2.58	2.589
16	2.169	2.387	2.766	2.654	2.654
17	2.052	2.289	2.745	2.67	2.764
18	2.056	2.342	2.8	2.482	2.669
19	2.064	2.284	2.874	2.52	2.511
20	2.165	2.216	2.849	2.53	2.717

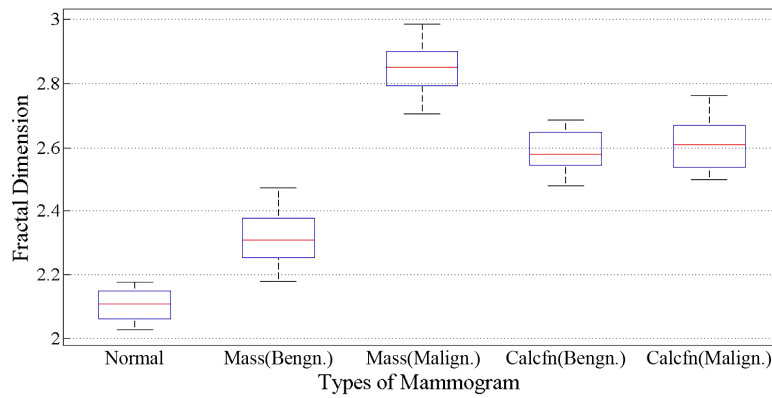
4.9.2.1. Box Plot of Fractal Dimensions

Box plot, also known as whisker plot is a convenient way of graphically depicting groups of numerical data. The main characteristics of this plot are the following:

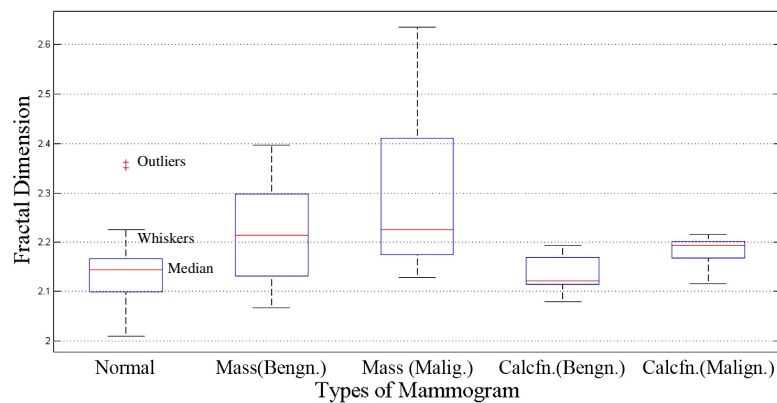
- The tops and bottoms of each "box" are the 25th and 75th percentiles of the samples, respectively. The distances between the tops and bottoms are the interquartile ranges.

- The line in the middle of each box is the sample median. If the median is not centered in the box, it shows sample skewness.
- The whiskers are lines extending above and below each box. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length (the adjacent values).
- Observations beyond the whisker length are marked as outliers. In the figure, outliers are displayed with a red + sign.
- Notches display the variability of the median between samples.

Box plots of fractal dimension obtained using different methods TPSA, Blanket and DBCM are shown in fig. 4 .6.



(a)TPSA



(b) DBCM

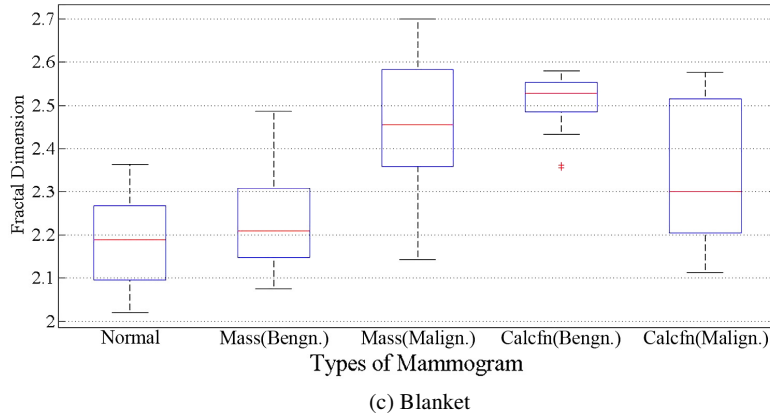


Fig. 4.6 Box Plot of the Fractal Dimension obtained using TPSA, DBCM and Blanket methods

It is clear from fig. 4.6(a), that for the TPSA method, the range of FD values for a few malignant masses and microcalcifications (both benign and malignant) are overlapping, so these categories cannot be correctly classified. In fig. 4.6(b), for DBCM, FD values of all the categories are overlapping each other, but the highest value was obtained for malignant mass and minimum value was obtained for the normal mammograms. Similar was the case with the blanket method shown in fig. 4.6(c).

4.9.2.2 Classification Accuracy using Fractal Dimension

The classification accuracy is the ratio of the number mammograms which are correctly classified to the total number of mammograms considered; both normal and cancerous. The table 4.7 shows the number of mammograms which are correctly classified in each database. In the TPSA method, the ranges of individual FD values were not overlapping for normal and benign and malignant masses. Therefore it was possible to correctly classify these categories with 100% accuracy.

But the range of FD values for benign and malignant microcalcifications were overlapping with the other categories. For the MIAS database, 166 Normal, 54

masses (benign) and 39 masses (malignant) were classified accurately, thus the classification accuracy becomes $259/287 = 90.24\%$.

Table 4.7

Mammograms correctly classified using fractal dimension computed using TPSA, DBC and Blanket methods

Database	No. of Mammograms correctly classified by Fractal Dimension Computed using						
	TPSA			DBC		Blanket	
	Normal	Mass (Ben)	Mass (Malig.)	Normal	Mass (Malig)	Normal	Mass (Malig)
MIAS (287 Nos)	166	54	39	13	9	14	8
DDSM (512 Nos)	180	84	75	21	14	21	14
Overall Classification accuracy%	74.84			7.1339		7.1339	

With the DDSM database, 180 normal, 84 masses (benign) and 75 masses (malignant) were correctly classified and the classification accuracy obtained was $339/512 = 66.21\%$. Thus the overall classification accuracy becomes 74.84%. As seen from table 4.5, in the DBC method, the range of FD values of various classes overlap each other. Only 9 malignant masses and 13 normal ones from the MIAS database and 14 malignant and 21 normal from the DDSM database were classified correctly using DBCM, giving a classification accuracy of 7.1339% only. Similar was the case with blanket method. Only 22 (14 normal and 8 malignant masses) from MIAS and 35 (21 normal and 14 malignant) from the DDSM databases could be accurately classified and the classification accuracy obtained was again 7.1339%.

It may be noted that, in the TPSA method, four experimental points considered at a time are forming a quadruple and this quadruple are covered by four triangles with mean elevation of four vertices as the common central point. When smaller triangular tiles are considered, they are not in simple relation with the cross section of the base of the prism, but also depend on the properties of the surface itself. Thus, TPSA method can provide an accurate measurement of fractal dimension

compared to the DBCM and blanket method. The latter two methods are similar, with the difference of the gray level surface is been considered for computing the fractal dimension.

As the classification accuracy for the feature FD is low as mentioned above, the six fractal features, $f_1 - f_6$ described in section 4.5 were calculated using the three FD estimation methods. They were found to provide better classification accuracy.

4.9.3 Evaluation using Fractal Features $f_1 - f_6$

The results obtained while computing the fractal features ($f_1 - f_6$) of malignant mass, benign mass, malignant and benign microcalcifications, and normal mammograms using the three methods are described in the following sections.

4.9.3.1 Evaluation using Fractal Feature f_1

Different overlapping window sizes were chosen to compute feature f_1 , but window of size of 2 gave good results. The feature images obtained using equation (4.24) to extract fractal feature f_1 is shown in fig. 4.7. The range of values, mean and standard deviation of f_1 computed using TPSA, DBCM and blanket method are shown in table 4.8. For the TPSA method, there was ample separation between the values for normal and cancerous mammograms, but the range of values for the classes overlap. For the other two methods ' f_1 ' values overlap between all classes.

The sample values of fractal feature f_1 obtained for 20 mammograms each from various classes, using the TPSA method, is presented in table 4.9.

The box plot for fractal feature f_1 obtained using TPSA; DBCM and Blanket methods are given in fig.4.8 (a)-(c) respectively. The results as analyzed in table 4.8 can be visualized in the box plot. The table 4.10 shows the number of mammograms which are correctly classified in each database. Table shows that none of the mammograms were classified correctly by DBC method as all the individual feature values were overlapping with other classes.

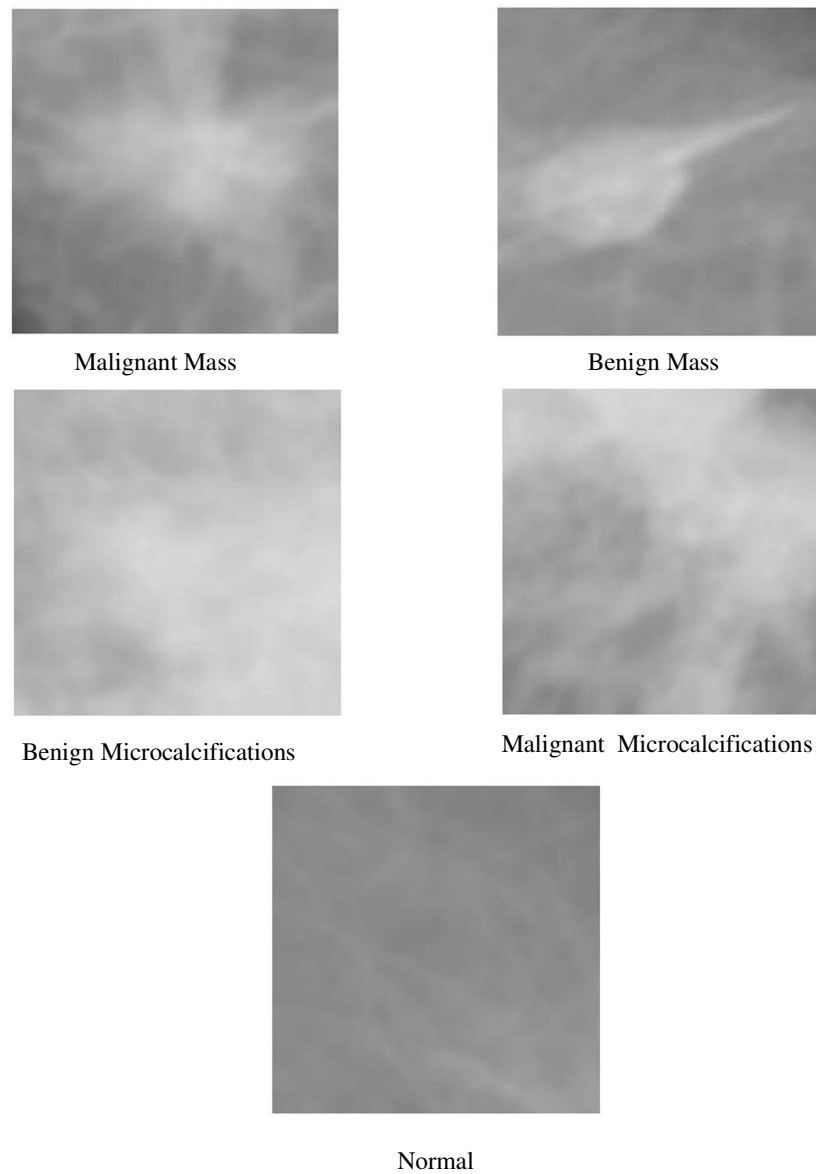


Fig. 4.7 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_1

In blanket method also, as seen from the box plot fig. 4.8(c) feature values of all the classes of mammograms were overlapping, giving an accuracy of only 1.877% as

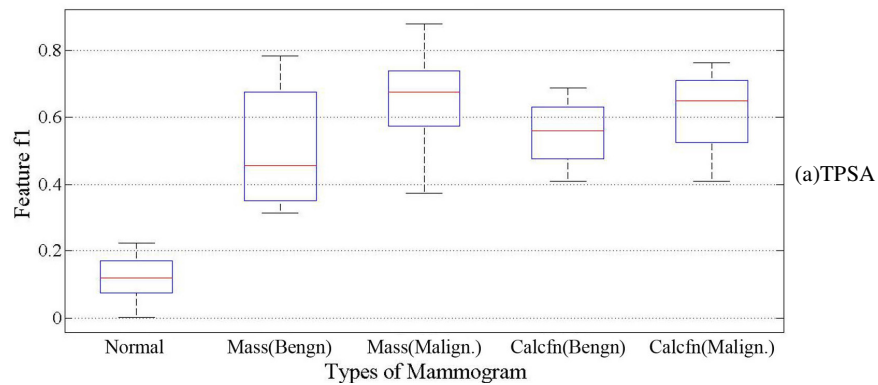
only 15 malignant masses were classified correctly. In the TPSA method, 296 out of 344 normal and 93 out of 252 masses were correctly detected giving a classification accuracy of 48.6858.

Table 4.8
Comparison of the fractal feature f_1 obtained using TPSA, DBC and Blanket Methods

		Fractal feature f_1 computed using								
		TPSA			DBC			Blanket		
Mammograms		Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev
Normal		0.002 - 0.223	0.1217	0.0668	0.153-0.451	0.2900	0.0863	0.0432-0.2963	0.1771	0.0746
Masses	Benign	0.312 - 0.784	0.5087	0.1621	0.046-0.299	0.1662	0.0892	0.0182-0.3445	0.1711	0.0845
	Malign.	0.372 - 0.879	0.658	0.1161	0.171-0.235	0.2110	0.0163	0.1611-0.3752	0.2682	0.0598
Microcalcifications	Benign	0.411 - 0.688	0.5578	0.0895	0.21-0.432	0.3288	0.0669	0.0787-0.3047	0.2392	0.0525
	Malign.	0.411 - 0.763	0.6201	0.1059	0.266-0.398	0.3264	0.0375	0.1692-0.3367	0.2624	0.0504
Classification Accuracy %		48.6858			0			1.877		

Table 4.9 Sample Fractal feature f_1 values obtained for different mammograms using TPSA method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	0.157	0.689	0.78	0.678	0.525
2	0.002	0.487	0.703	0.501	0.457
3	0.115	0.312	0.568	0.566	0.638
4	0.099	0.45	0.372	0.452	0.712
5	0.119	0.346	0.781	0.517	0.411
6	0.067	0.55	0.703	0.533	0.653
7	0.102	0.559	0.6763	0.607	0.689
8	0.102	0.335	0.695	0.632	0.608
9	0.095	0.378	0.527	0.599	0.66
10	0.054	0.356	0.622	0.683	0.601
11	0.168	0.452	0.604	0.663	0.644
12	0.07	0.528	0.534	0.632	0.73
13	0.19	0.436	0.738	0.6045	0.696
14	0.19	0.337	0.632	0.5744	0.707
15	0.002	0.369	0.671	0.411	0.6247
16	0.053	0.457	0.748	0.523	0.711
17	0.15	0.784	0.538	0.681	0.732
18	0.127	0.599	0.392	0.532	0.634
19	0.143	0.641	0.592	0.59	0.677
20	0.17	0.45	0.709	0.621	0.431



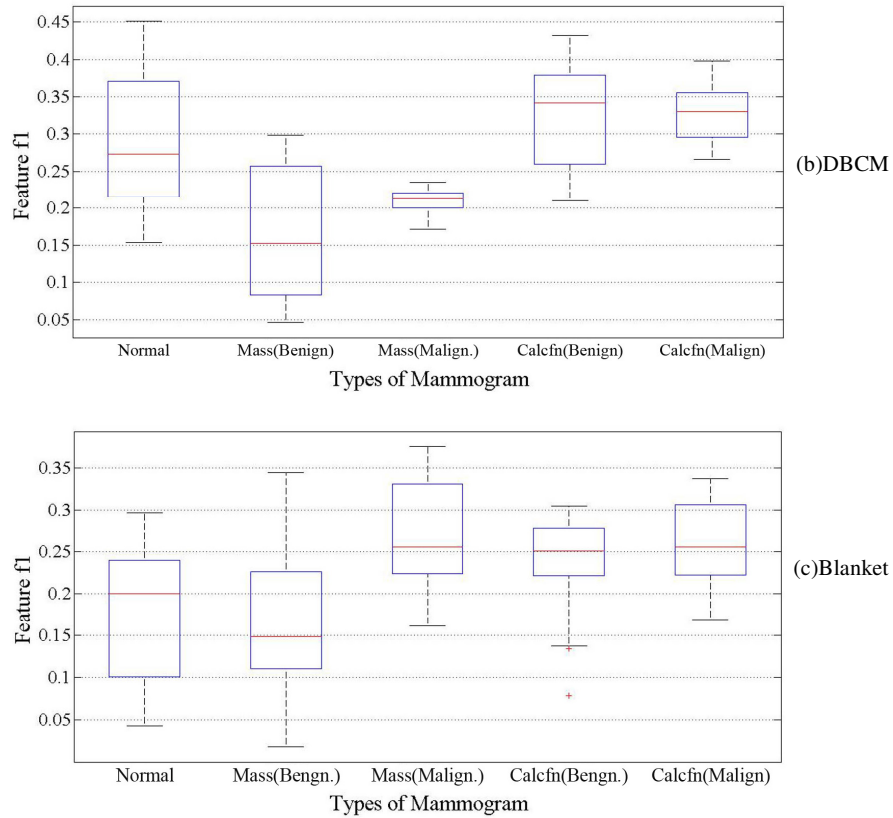


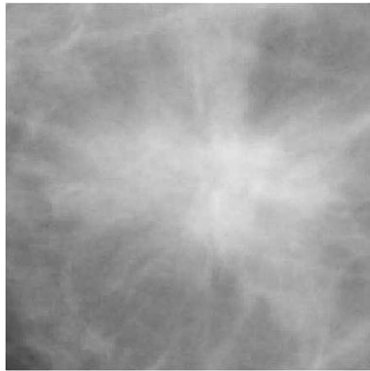
Fig. 4.8 Box Plot of the feature f_1 obtained using TPSA, DBCM and Blanket methods

Table 4.10
Mammograms Classification using Fractal feature f_1 computed using TPSA, DBC and Blanket methods

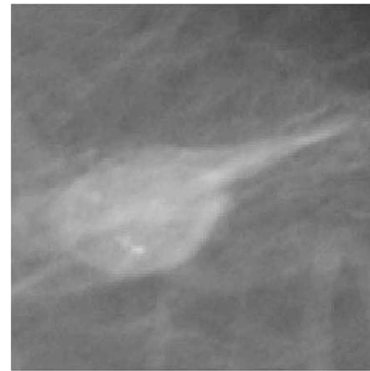
Database	No. of Mammograms Classified Correctly by Fractal feature f_1 Computed using				
	Normal	TPSA Mass (Ben)	Mass (Malig.)	DBC -	Blanket Mass (Malig)
MIAS(287 Nos)	126	15	11	-	5
DDSM(512 Nos)	170	36	31	-	10
Overall Classification accuracy %		48.6858		0	1.877

4.9.3.2 Evaluation using Fractal Feature f_2

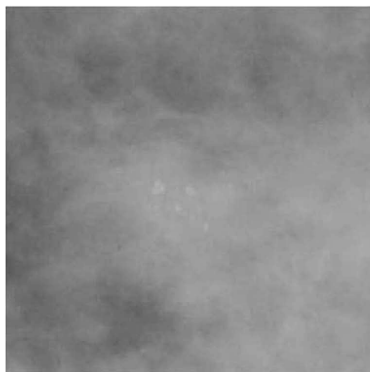
The high gray valued images obtained using equation (4.25), for the different categories of mammograms are shown in fig.4.9. Table 4.11 illustrates the comparison of the fractal feature f_2 values obtained for different methods. In the TPSA method, the maximum value for feature f_2 was obtained for the malignant mammograms while minimum was obtained for the normal mammogram, as required. Sample feature f_2 values obtained by the TPSA method are shown in table 4.12. The spread of the fractal feature f_2 values are illustrated in the box plot fig. 4.10. As seen from box plot 4.10(a), none of the individual values of feature f_2 for normal mammograms were overlapping with other classes, but all the other classes had overlapping values. Therefore, all normal mammograms and few masses were correctly classified by feature f_2 , obtained by TPSA method.



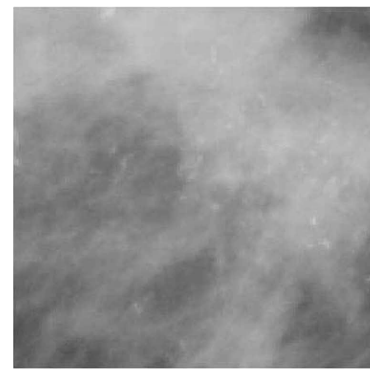
Malignant Mass



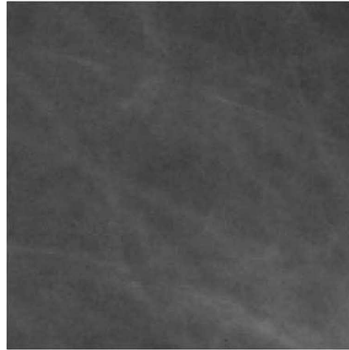
Benign Mass



Benign Microcalcifications



Malignant Microcalcifications



Normal

Fig.4.9 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_2

For the DBC method, benign masses had the lowest value for the feature f_2 , but this cannot be considered as correct evaluation as benign mammograms have more irregularity than normal ones. Therefore, f_2 should be higher for benign than for normal. Only 27 (7 MIAS and 20 DDSM) malignant masses were classified correctly and the overall classification accuracy was only 3.379%.

Table 4.11
Comparison of the fractal feature f_2 obtained using TPSA, DBC and Blanket
Methods

Mammograms		Fractal feature f_2 computed using								
		TPSA			DBC			Blanket		
		Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev
Normal		0.002-0.223	0.1201	0.0788	0.153-0.451	0.296	0.0843	0.043-0.296	0.190	0.1
Masses	Benign	0.312-0.784	0.4962	0.1371	0.046-0.299	0.1707	0.0703	0.018-0.345	0.218	0.15
	Malignant	0.524-0.879	0.6465	0.5585	0.171-0.635	0.3788	0.1397	0.161-0.375	0.264	0.069
Microcalcifications	Benign	0.411-0.688	0.5585	0.0776	0.21-0.432	0.3274	0.0620	0.221-0.339	0.290	0.183
	Malignant	0.501-0.763	0.6197	0.0698	0.266-0.397	0.3426	0.0355	0.221-0.361	0.263	0.073
Classification Accuracy %		46.433			3.379			0.6257		

Table 4.12 Sample Fractal feature f_2 obtained for different mammograms using TPSA method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	0.157	0.682	0.751	0.562	0.62
2	0.02	0.347	0.685	0.411	0.54
3	0.115	0.312	0.524	0.459	0.6351
4	0.002	0.419	0.752	0.56	0.571
5	0.002	0.391	0.762	0.43	0.518
6	0.067	0.507	0.661	0.486	0.5194
7	0.102	0.534	0.605	0.489	0.523
8	0.2102	0.39	0.731	0.644	0.6009
9	0.102	0.335	0.557	0.584	0.687
10	0.095	0.338	0.638	0.606	0.592
11	0.02	0.535	0.574	0.5742	0.5446
12	0.2002	0.476	0.545	0.507	0.6105
13	0.202	0.523	0.636	0.4275	0.7105
14	0.031	0.389	0.643	0.411	0.604
15	0.002	0.451	0.666	0.688	0.585
16	0.129	0.402	0.555	0.534	0.5433
17	0.129	0.756	0.653	0.581	0.501
18	0.128	0.67	0.538	0.573	0.763
19	0.1009	0.541	0.681	0.653	0.665
20	0.1039	0.37	0.735	0.591	0.649

For the blanket method, only five (2 MIAS and 3 DDSM) malignant masses, which had fractal feature f_2 which were not overlapping with other classes, and could be correctly classified and the classification accuracy was found to be 0.6257%.

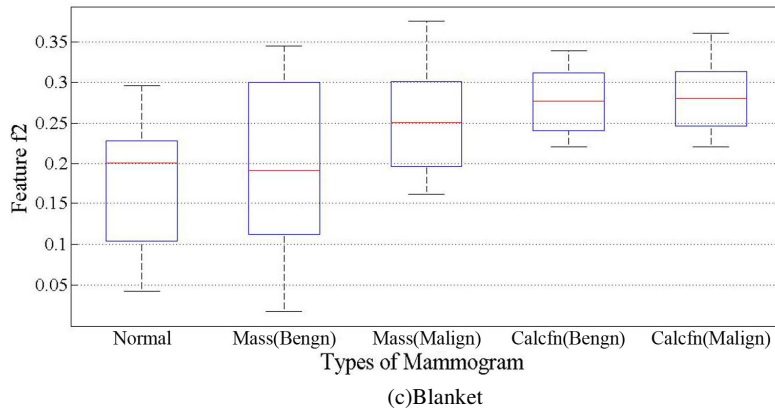
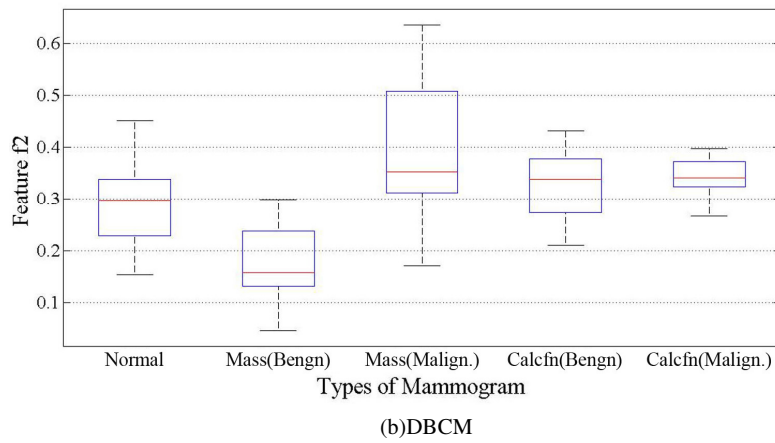
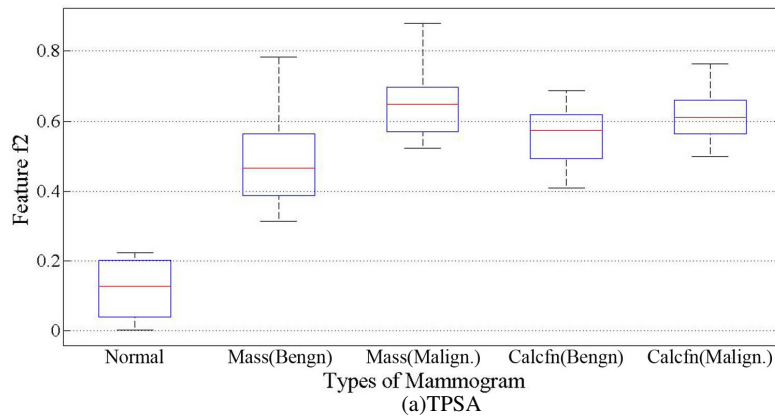


Fig. 4.10 Box Plot of the feature f_2 obtained using TPSA, DBCM and Blanket methods

The actual number of mammograms which were correctly classified in each database, using the different methods is shown in table 4.13.

Table 4.13
Mammograms correctly classified using Fractal feature f_2 computed using TPSA, DBC and Blanket methods

Database	No. of Mammograms Classified Correctly by Fractal feature f_2 Computed using				
	TPSA		DBC		Blanket
	Normal	Mass (Ben)	Mass (Malig.)	Mass (Malig)	Mass (Malig)
MIAS (287 Nos)	166	7	4	7	2
DDSM (512 Nos)	180	10	4	20	3
Overall Classification accuracy%		46.433		3.379	0.6257

4.9.3.3 Evaluation using Fractal Feature f_3

The low gray level feature images computed using equation (4.26) for finding fractal feature f_3 is shown in fig.4.11. As seen in fig.4.11, no structures of the mammograms are present in the f_3 feature image and hence no details about the structure will be obtained from this feature f_3 . This is evident from the values shown in table 4.14, which gives the comparison of f_3 feature values obtained using the three methods. Table 4.14 shows that the values of the fractal feature f_3 , are in the same range for the different classes of mammograms, for all the three methods and these values are found to be very small. This is because irregularity in the low gray level image is less. The sample values of fractal feature f_3 obtained for the different classes of mammo-

grams using the TPSA method is presented in table 4.15. The box plot given in fig.4.12 also indicates that all the categories of mammograms are having values in the same range.

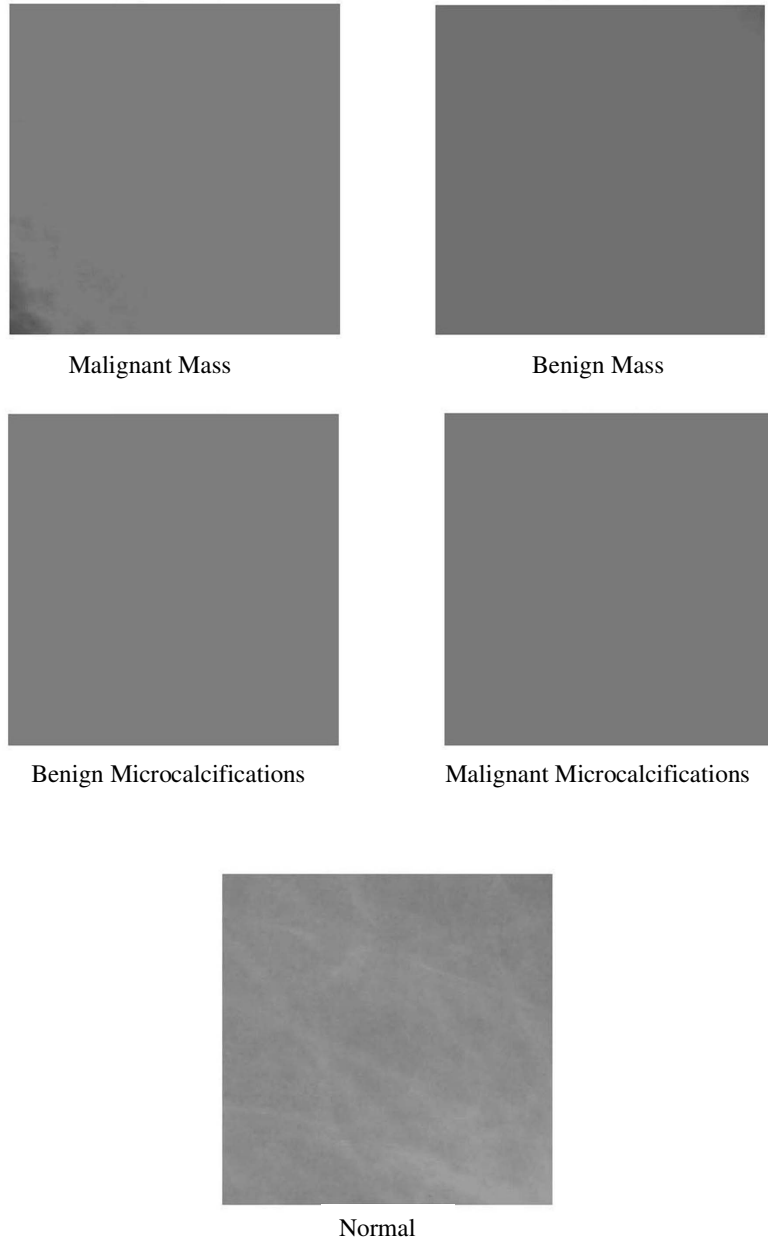


Fig.4.11 Fractal feature image of Malignant mass, benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_3

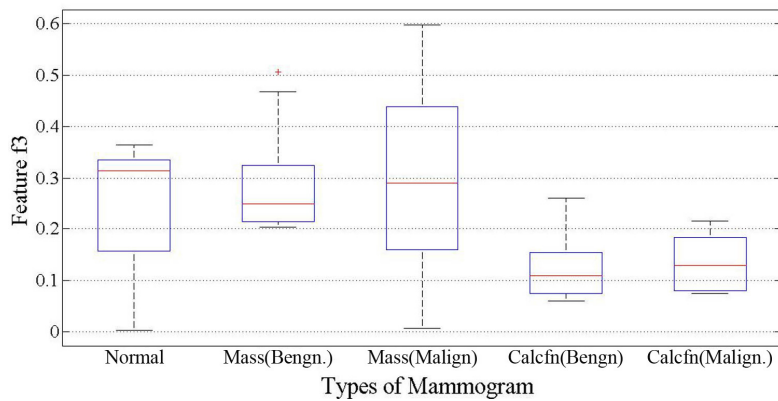
The least classification accuracy was obtained for this feature, with the values being 1.5018%, 0% and 1.5018% for the TPSA, DBC and Blanket methods respectively. Table 4.16 shows the number of mammograms which are correctly classified in each database using the fractal feature f_3 .

Table 4.14
Comparison of the fractal feature f_3 obtained using TPSA, DBC and Blanket Methods

Mammo grams		Fractal feature f_3 computed using								
		TPSA			DBC			Blanket		
		Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev
Normal		0.003-0.365	0.2532	0.1096	0.144- 0.531	0.347	0.1297	0-0.012	0.0071	0.0037
Masses	Benign	0.203-0.506	0.2785	0.0841	0.074- 0.368	0.1834	0.0975	0-0.12	0.01019	0.0187
	Mali gnant	0.007-0.598	0.2894	0.1515	0.077- 0.492	0.244	0.115	0-0.1127	0.036	0.0458
Micro calcific ations	Benign	0.06-0.262	0.125	0.065	0.074- 0.251	0.184	0.0555	0-0.003	0.00142	0.00109
	Mali gnant	0.075-0.215	0.1324	0.0503	0.042- 0.144	0.107	0.0294	0-0.007	0.0045	0.0022
Classification Accuracy %		1.5018			0			1.5018		

Table 4.15 Sample Fractal feature f_3 obtained for different mammograms using TPSA method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	0.315	0.468	0.138	0.262	0.215
2	0.365	0.203	0.497	0.075	0.075
3	0.315	0.21	0.215	0.075	0.211
4	0.365	0.207	0.215	0.0648	0.075
5	0.337	0.248	0.091	0.0754	0.193
6	0.253	0.203	0.453	0.067	0.075
7	0.2521	0.263	0.435	0.075	0.075
8	0.365	0.235	0.089	0.132	0.076
9	0.365	0.259	0.131	0.075	0.205
10	0.14	0.506	0.375	0.075	0.126
11	0.259	0.417	0.373	0.165	0.129
12	0.365	0.221	0.44	0.215	0.089
13	0.365	0.361	0.007	0.075	0.077
14	0.314	0.252	0.429	0.155	0.075
15	0.129	0.205	0.464	0.06	0.19
16	0.003	0.218	0.172	0.11	0.086
17	0.123	0.286	0.23	0.13	0.0944
18	0.365	0.289	0.215	0.124	0.104
19	0.221	0.241	0.159	0.09	0.15
20	0.2201	0.2065	0.133	0.091	0.201



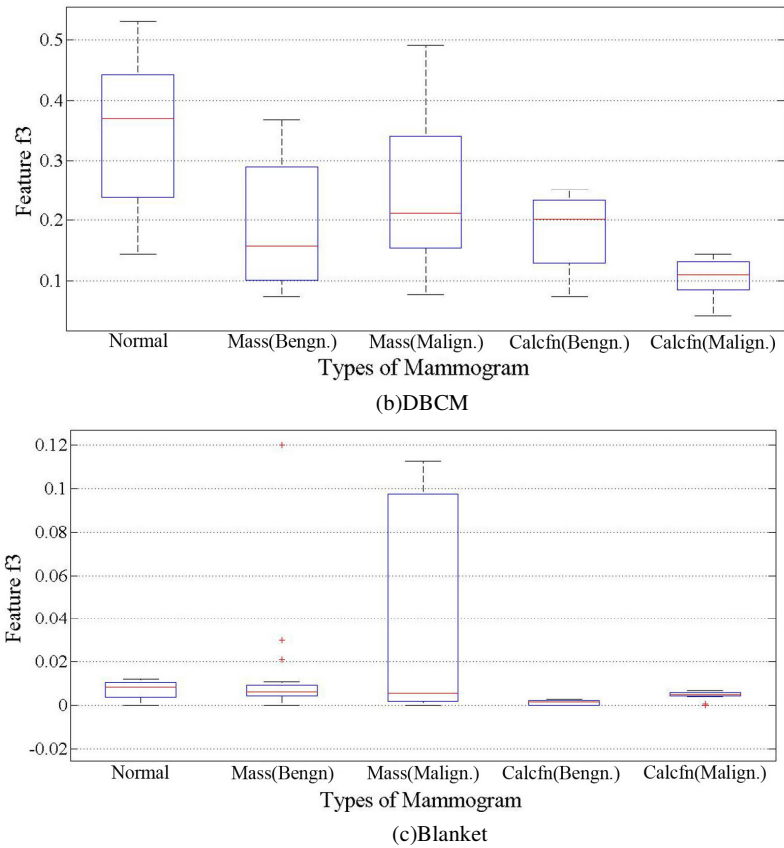


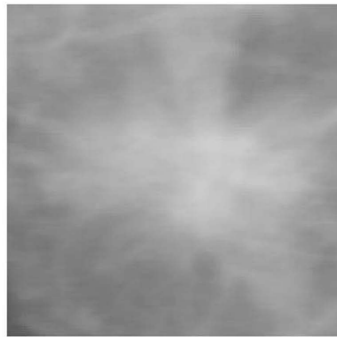
Fig 4.12 Box Plot of the feature f_3 obtained using TPSA, DBCM and Blanket methods

Table 4.16
Mammograms correctly classified using Fractal feature f_3 computed using TPSA, DBC and Blanket methods

Database	No. of Mammograms Classified Correctly by Fractal feature f_3 Computed using			
	TPSA	DBC	Blanket	
	Normal	Mass (Malig.)	-	Mass (Malig)
MIAS (287 Nos)	1	3	0	4
DDSM (512 Nos)	2	6	0	8
Overall Classification accuracy%	1.5018		0	1.5018

4.9.3.4 Evaluation using Fractal Feature f_4

The horizontally smoothed images computed using equation (4.28) for finding the fractal feature f_4 is given in fig.4.13. Value of feature f_4 obtained using the three different methods are given in tables 4.17. The box plot of the feature values for the TPSA, DBC and Blanket methods are given in fig.4.14. The normal mammograms and malignant masses had the lowest and highest feature f_4 values respectively for the TPSA and DBC methods. But the standard deviation of this feature for the malignant mass class were obtained as 0.105 and 0.1496 respectively for the above two methods causing overlap between the classes. For the blanket method, the values of the fractal feature f_4 were in the same range for all the classes of mammograms.



Malignant Mass



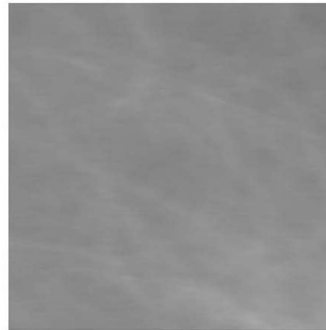
Benign Mass



Benign Microcalcifications



Malignant Microcalcifications



Normal

Fig.4.13 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_4

As seen from table 4.17, the highest classification accuracy obtained for the feature f_4 was using the DBC method. Therefore, 20 sample values of fractal feature f_4 , obtained for different mammograms using DBC method are shown in table 4.18.

As shown in table 4.19, in TPSA, 26 normal (11 from MIAS and 15 from DDSM) and 12 malignant masses (4 MIAS and 8 DDSM) were classified correctly.

Table 4.17
Comparison of the fractal feature f_4 obtained using TPSA, DBC and Blanket Methods

Fractal feature f_4 computed using										
Mammograms	TPSA			DBC			Blanket			
	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	
Normal	0.018-0.173	0.0806	0.0449	0.055-0.582	0.2583	0.1297	0.0497-0.4771	0.325	0.1325	
Masses	Benign	0.167-0.681	0.3512	0.1258	0.218-0.44	0.332	0.0693	0.1298-0.5658	0.3375	0.1375
	Malignant	0.337-0.72	0.4964	0.105	0.242-0.758	0.4667	0.1496	0.1039-0.4391	0.2598	0.1041
Micro calcifications	Benign	0.14-0.696	0.2566	0.1132	0.407-0.517	0.450	0.0384	0.3979-0.5919	0.481	0.0583
	Malignant	0.068-0.378	0.1881	0.0932	0.251-0.477	0.4049	0.0625	0.0822-0.4035	0.3036	0.0914
Classification Accuracy %	4.755			16.2703			2.377			

Highest mean value for the feature value was observed for malignant and the least for normal mammograms. Even if there is enough separation between the mean values of the different classes the individual values were spread between the classes as observed in the box plot fig.4.14 (a).

With the DBC method, more number of normal mammograms (32 MIAS and 51 DDSM) and malignant masses (13- MIAS and 34 DDSM) were correctly classified. Thus the overall classification accuracy with this feature was found to be 16.2703%. For this feature alone, DBCM gave the highest classification accuracy.

Table 4.18 Sample Fractal feature f_4 obtained for different mammograms using DBC method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	0.273	0.367	0.625	0.456	0.451
2	0.185	0.313	0.512	0.591	0.512
3	0.1009	0.329	0.765	0.539	0.5227
4	0.0118	0.3	0.656	0.532	0.47
5	0.213	0.344	0.59	0.547	0.3456
6	0.217	0.204	0.644	0.345	0.5119
7	0.235	0.298	0.51	0.3556	0.539
8	0.022	0.318	0.54	0.3689	0.3557
9	0.0213	0.272	0.759	0.4878	0.3497
10	0.104	0.3176	0.501	0.4876	0.3965
11	0.249	0.2954	0.746	0.4556	0.4853
12	0.345	0.2134	0.5667	0.4431	0.4467
13	0.344	0.2451	0.663	0.3782	0.4838
14	0.0321	0.3671	0.6689	0.3679	0.4253
15	0.361	0.3976	0.5691	0.3115	0.4735
16	0.0125	0.3566	0.7345	0.3968	0.4863
17	0.2781	0.3588	0.7293	0.4856	0.4291
18	0.121	0.3744	0.5792	0.4552	0.5003
19	0.125	0.408	0.5631	0.4483	0.5047
20	0.3712	0.3785	0.5741	0.5341	0.4931

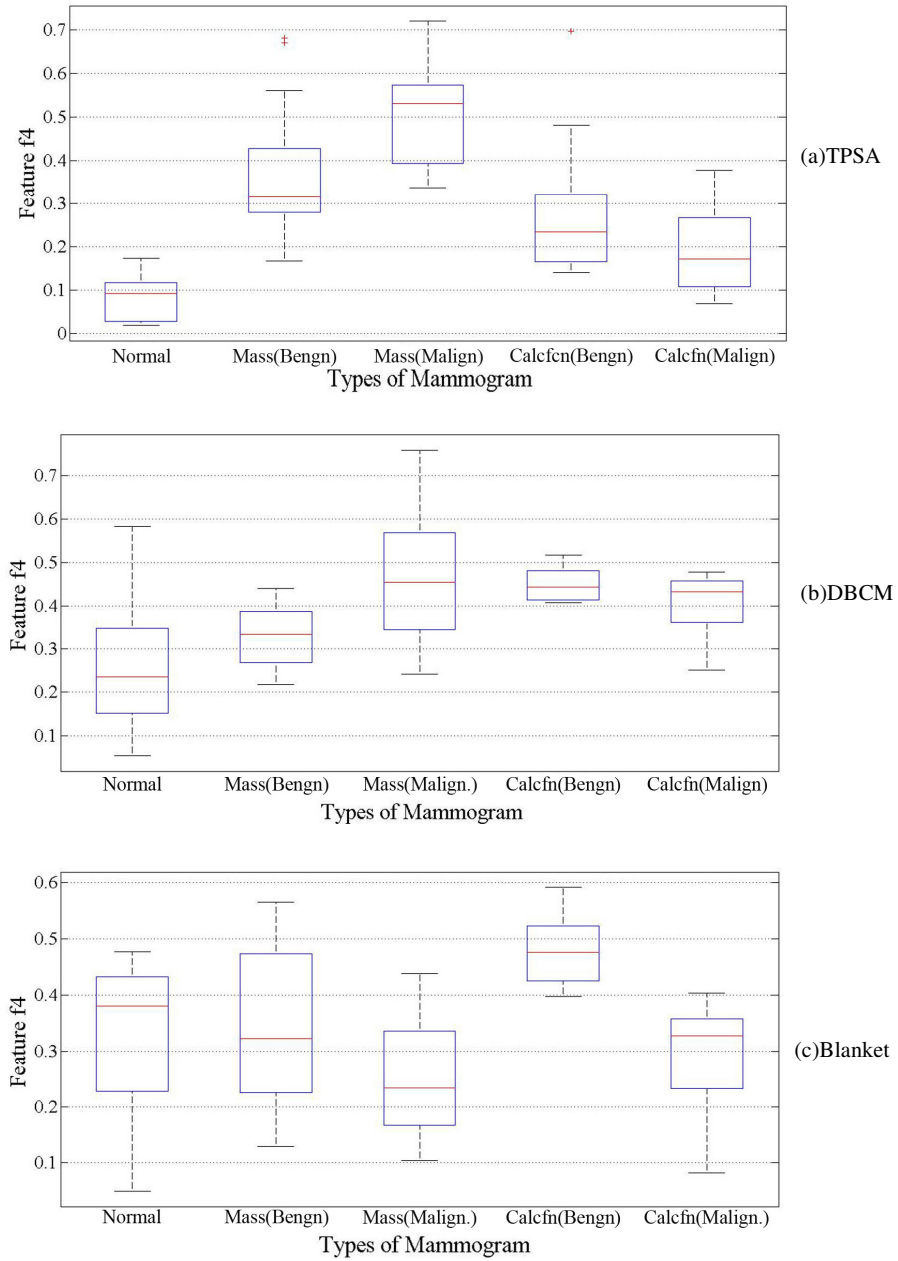


Fig. 4.14 Box Plot of the feature f_4 obtained using TPSA, DBCM and Blanket methods

Table 4.19
Mammograms correctly classified using Fractal feature f_4 computed using TPSA, DBC and Blanket methods

Database	No. of Mammograms Classified Correctly by Fractal feature f_4 Computed using				
	TPSA		DBC		Blanket
	Normal	Mass (Malig.)	Normal	Mass (Malig.)	Normal
MIAS (287 Nos)	11	4	32	13	5
DDSM (512 Nos)	15	8	51	34	14
Overall Classification accuracy%	4.755		16.2703		2.377

In the blanket method, the benign microcalcifications produced the highest feature values. Also normal and malignant microcalcifications had the least values which are not correct according to the argument. Thus only 19 normal (5 MIAS and 14 DDSM) mammograms were classified correctly with a classification accuracy of 2.377%.

4.9.3.5 Evaluation using Fractal Feature f_5

Fig.4.15 shows the vertically smoothed image obtained using equation (4.29) for computing the fractal feature f_5 , for the different classes of mammograms. The feature values obtained for the three methods are tabulated in table 4.20. 20 sample values of fractal feature f_5 obtained for the different classes of mammograms using the TPSA method is presented in table 4.21.

The spread of the feature values f_5 is illustrated in the box plot shown in fig. 4.16. Normal and malignant masses had the lowest and highest feature f_5 values respectively for TPSA and DBC method. But these values were found to be overlapping between other classes. In blanket method, the values are spread over all the classes and could not distinguish between any class of mammograms as shown in fig.4.16(a)-(c).The mean values also of normal, benign mass and malignant microcalcifications were found to be in the same range for the blanket method.

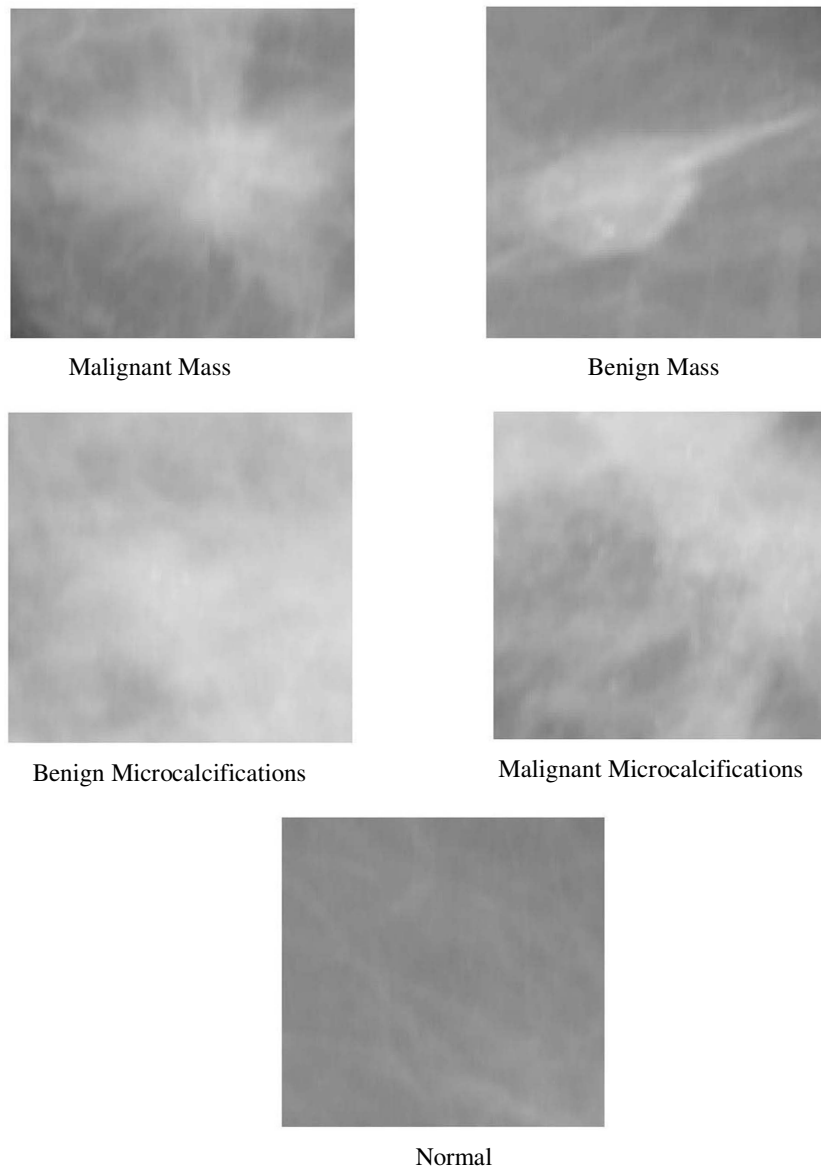


Fig.4.15 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_5

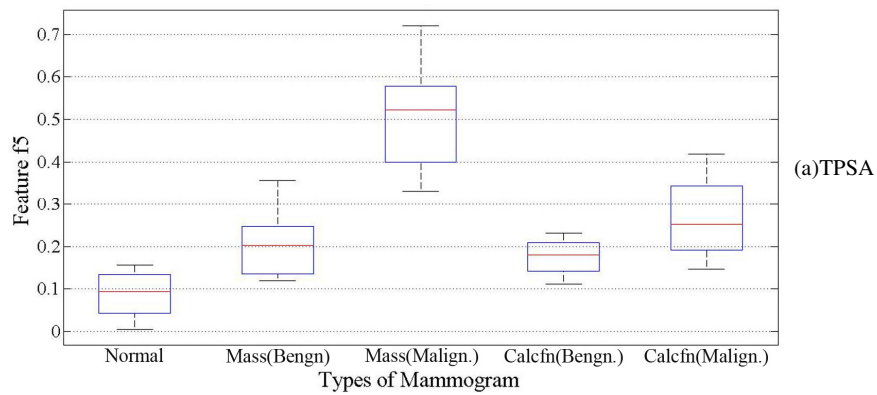
Table 4.20
Comparison of the fractal feature f_5 obtained using TPSA, DBC and Blanket Methods

Mammo grams		Fractal feature f_5 computed using								
		TPSA			DBC			Blanket		
		Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev
Normal		0.005- 0.156	0.0893	0.0497	0.113- 0.342	0.237	0.0595	0.0603- 0.5004	0.3475	0.1298
Masses	Benign	0.119- 0.357	0.2087	0.076	0.204- 0.421	0.3342	0.0637	0.1025- 0.4469	0.3173	0.089
	Mali gnant	0.331- 0.72	0.4993	0.1053	0.194- 0.727	0.4886	0.1218	0.0658- 0.4485	0.2809	0.1094
Micro calcific ations	Benign	0.111- 0.231	0.1777	0.0379	0.266- 0.493	0.3873	0.0593	0.3944- 0.567	0.4699	0.0450
	Mali gnant	0.146- 0.419	0.2681	0.0856	0.293- 0.524	0.4188	0.0647	0.1925- 0.4628	0.3484	0.0729
Classification Accuracy %		36.5456			19.0237			0		

The number of mammograms which are correctly classified using feature f_5 is presented in table 4.22. Classification accuracies of 36.5456% (MIAS-normal (81), malignant mass (24), DDSM-normal (115), malignant mass (72)), 19.0237% (MIAS-normal (17), malignant mass (25), DDSM-normal (41), malignant mass (69)) and 0% were obtained for TPSA, DBCM and Blanket methods respectively

Table 4.21 Sample Fractal feature f_5 obtained for different mammograms using TPSA method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	0.119	0.241	0.578	0.127	0.146
2	0.005	0.121	0.595	0.125	0.191
3	0.118	0.121	0.383	0.111	0.219
4	0.0067	0.119	0.331	0.16	0.383
5	0.0067	0.123	0.561	0.169	0.235
6	0.0067	0.1478	0.535	0.2	0.146
7	0.09	0.2022	0.523	0.221	0.396
8	0.079	0.274	0.525	0.141	0.157
9	0.093	0.141	0.336	0.14	0.306
10	0.094	0.119	0.564	0.175	0.281
11	0.094	0.3514	0.497	0.22	0.1533
12	0.094	0.357	0.48	0.231	0.1689
13	0.094	0.188	0.453	0.138	0.419
14	0.0052	0.128	0.585	0.1925	0.186
15	0.0052	0.301	0.595	0.1807	0.1814
16	0.126	0.215	0.437	0.1922	0.345
17	0.145	0.348	0.384	0.1895	0.229
18	0.145	0.241	0.341	0.2134	0.3195
19	0.018	0.23	0.349	0.2094	0.381
20	0.032	0.247	0.52	0.2076	0.3445



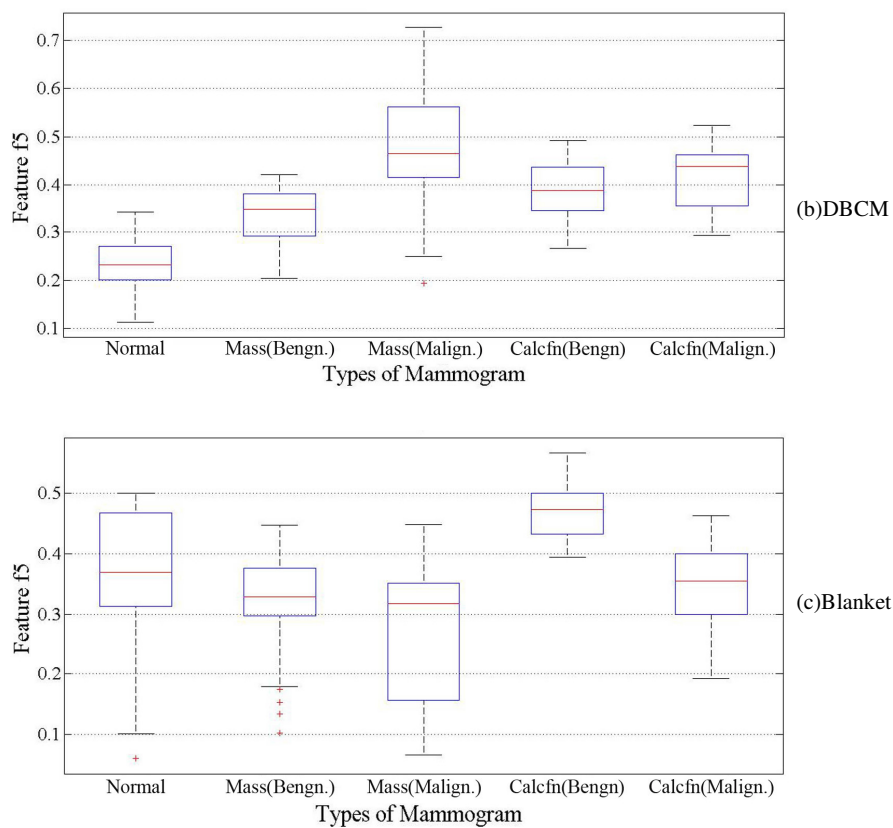


Fig.4.16 Box Plot of the feature f_5 obtained using TPSA, DBCM and Blanket methods

Table 4.22
Mammograms correctly classified using Fractal feature f_5 computed using TPSA, DBC and Blanket methods

Database	No. of Mammograms Classified Correctly by Fractal feature f_5 Computed using				
	TPSA		DBC		Blanket
	Normal	Mass (Malig.)	Normal	Mass (Malig.)	-
MIAS (287 Nos)	81	24	17	25	-
DDSM (512 Nos)	115	72	41	69	-
Overall Classification accuracy %	36.5456		19.0237		-

4.9.3.6 Evaluation using Fractal Feature f_6

As seen from the results in sections 4.9.3.1 to 4.9.3.5 none of the existing fractal features $f_1 - f_5$, were able to classify mammograms effectively. Hence the new fractal feature f_6 was developed from the average of four pixels as mentioned in equation (4.30). The evaluation of f_6 was done using TPSA, DBC and Blanket methods. It was found that, rather than finding the difference between the minimum and maximum gray level, in the DBCM method, and finding the volume of the blanket in the blanket method, the area of the triangles formed by the height difference in the gray levels gave better classification accuracy than the other two methods.

The images obtained for the different classes of mammograms, used for calculating the fractal feature f_6 are shown in fig.4.17. Table 4.23 illustrates the feature values obtained by different methods. It is seen from the table that this feature could classify the different types of mammograms more effectively and efficiently for the TPSA method.

A set of 20 sample values of fractal feature f_6 , obtained for the different classes of mammograms using the TPSA method is presented in table 4.24. The spread of the fractal feature f_6 as seen in the box plot of fig.4.18(a)-(c) also projects the same result. Analyzing the results for the TPSA method, it was seen that, there was large separation between the f_6 values and 100% classification was possible with the 287 mammograms in MIAS database. For the 512 mammograms

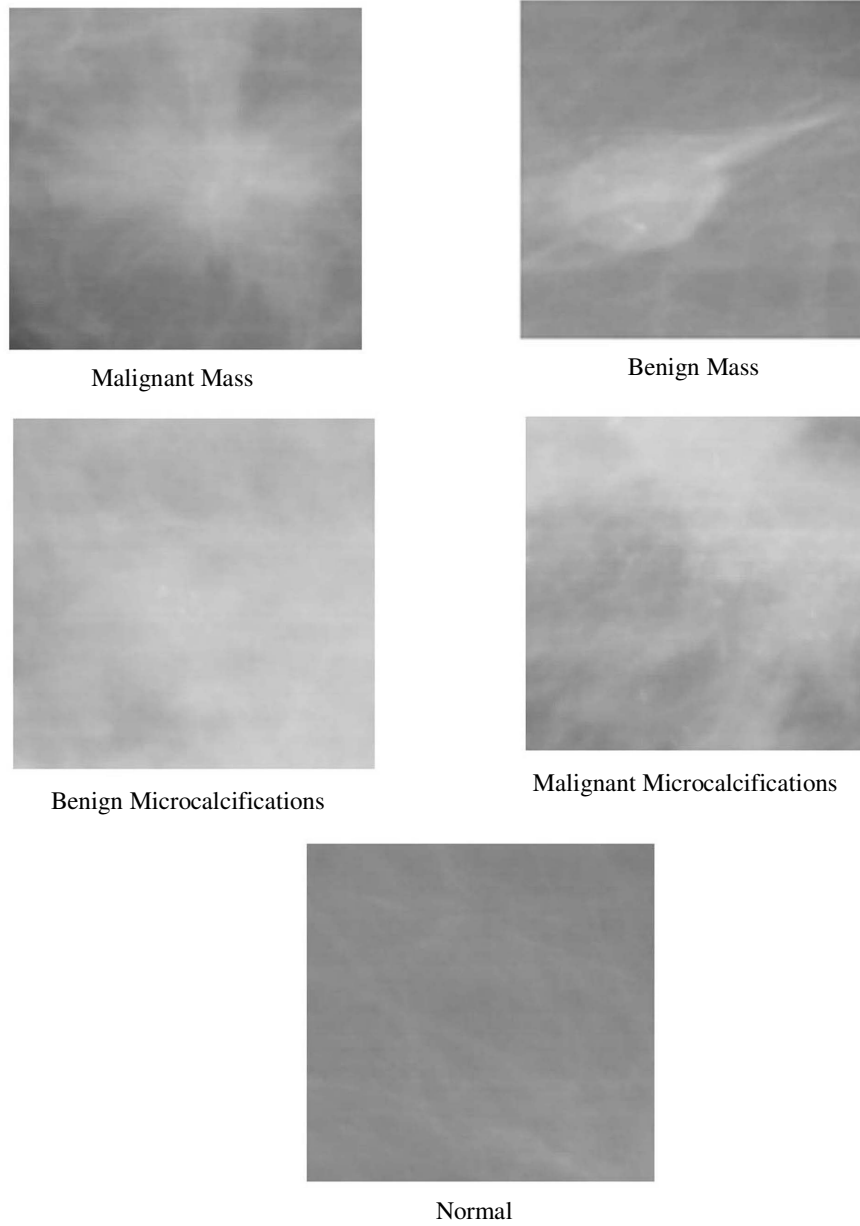


Fig.4.17 Fractal feature image of Malignant Mass, Benign Mass, Benign Microcalcifications, Malignant Microcalcifications, Normal Mammograms respectively for computing feature f_6

in the DDSM database, there was an overlap between the values 0.401-0.628 for 11 benign and malignant microcalcifications. In the case of the remaining mammograms, for normal mammograms, the feature f_6 was found to be between 0.001 and 0.156, for benign microcalcifications 0.401 to 0.601, malignant microcalcifications 0.56 to 0.62, benign masses 0.209 to 0.35 and malignant masses 0.7 to 0.986. The classification accuracy of the DDSM database was found to be $501/512 = 97.85\%$. Thus the overall classification accuracy of this feature was found to be 98.623%. The box plot in fig.4.18 (a) shows the spread of values for the MIAS database. None of the individual values overlap and 100% classification was obtained in this case. Fig 4.18 (b) shows the box plot for the two databases together, wherein the small overlap, as mentioned earlier, for the 11 specific cases can be noted.

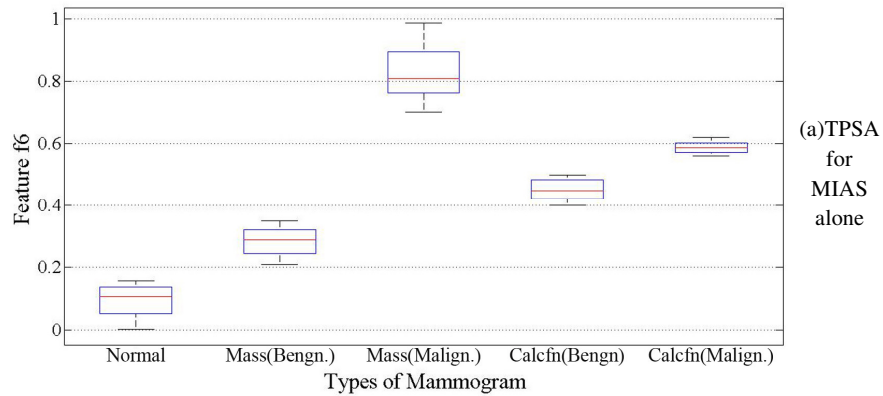
Fig. 4.18(c) and (d) show the values obtained for DBC and Blanket methods. For DBC method the individual values for the different classes are overlapping, but the mean values are different for different classes of mammograms. Also the difference between the mean values of benign and malignant microcalcifications was very less. In the MIAS database 66 normal and 23 malignant masses and in the DDSM database, 96 normal and 58 malignant masses were classified correctly using the fractal feature f_6 , with an overall classification accuracy of 30.413%.

Table 4.23Comparison of the fractal feature f_6 obtained using TPSA, DBC and Blanket Methods

Mammograms	Fractal feature f_6 computed using									
	TPSA			DBC			Blanket			
	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	Range of values	Mean	Std Dev	
Normal	0.001-0.156	0.0934	0.0485	0.0-0.388	0.1783	0.1221	0.0508-0.3646	0.1773	0.1232	
Benign	0.209-0.35	0.2841	0.0471	0.204-0.408	0.3142	0.0549	0.1029-0.4471	0.2078	0.1036	
Masses	Malignant	0.7-0.986	0.8269	0.0785	0.501-0.765	0.6178	0.0796	0.126-0.5258	0.3777	0.6612
Benign	0.401-0.601	0.4497	0.0327	0.312-0.591	0.4454	0.0807	0.1603-0.5778	0.2930	0.1731	
Microcalcifications	Malignant	0.56-0.62	0.5861	0.0171	0.3456-0.539	0.4516	0.0535	0.1678-0.2918	0.2311	0.0616
Classification Accuracy %	98.623			30.413			23.905			

Table 4.24 Sample Fractal feature f_6 obtained for different mammograms using TPSA method

Sl. No	Normal	Masses		Microcalcifications	
		Benign	Malignant	Benign	Malignant
1	0.097	0.318	0.764	0.484	0.5753
2	0.089	0.327	0.753	0.5	0.578
3	0.14	0.209	0.762	0.402	0.5822
4	0.025	0.273	0.986	0.424	0.577
5	0.136	0.259	0.863	0.409	0.56
6	0.104	0.298	0.824	0.435	0.583
7	0.134	0.22	0.798	0.461	0.592
8	0.128	0.306	0.791	0.413	0.563
9	0.139	0.337	0.745	0.409	0.562
10	0.156	0.3412	0.779	0.422	0.562
11	0.107	0.321	0.823	0.433	0.5782
12	0.024	0.239	0.949	0.479	0.581
13	0.124	0.223	0.886	0.489	0.561
14	0.132	0.283	0.941	0.422	0.564
15	0.052	0.219	0.849	0.409	0.59
16	0.001	0.285	0.803	0.421	0.568
17	0.052	0.285	0.73	0.411	0.567
18	0.056	0.297	0.923	0.433	0.62
19	0.064	0.244	0.876	0.401	0.601
20	0.005	0.291	0.749	0.486	0.602



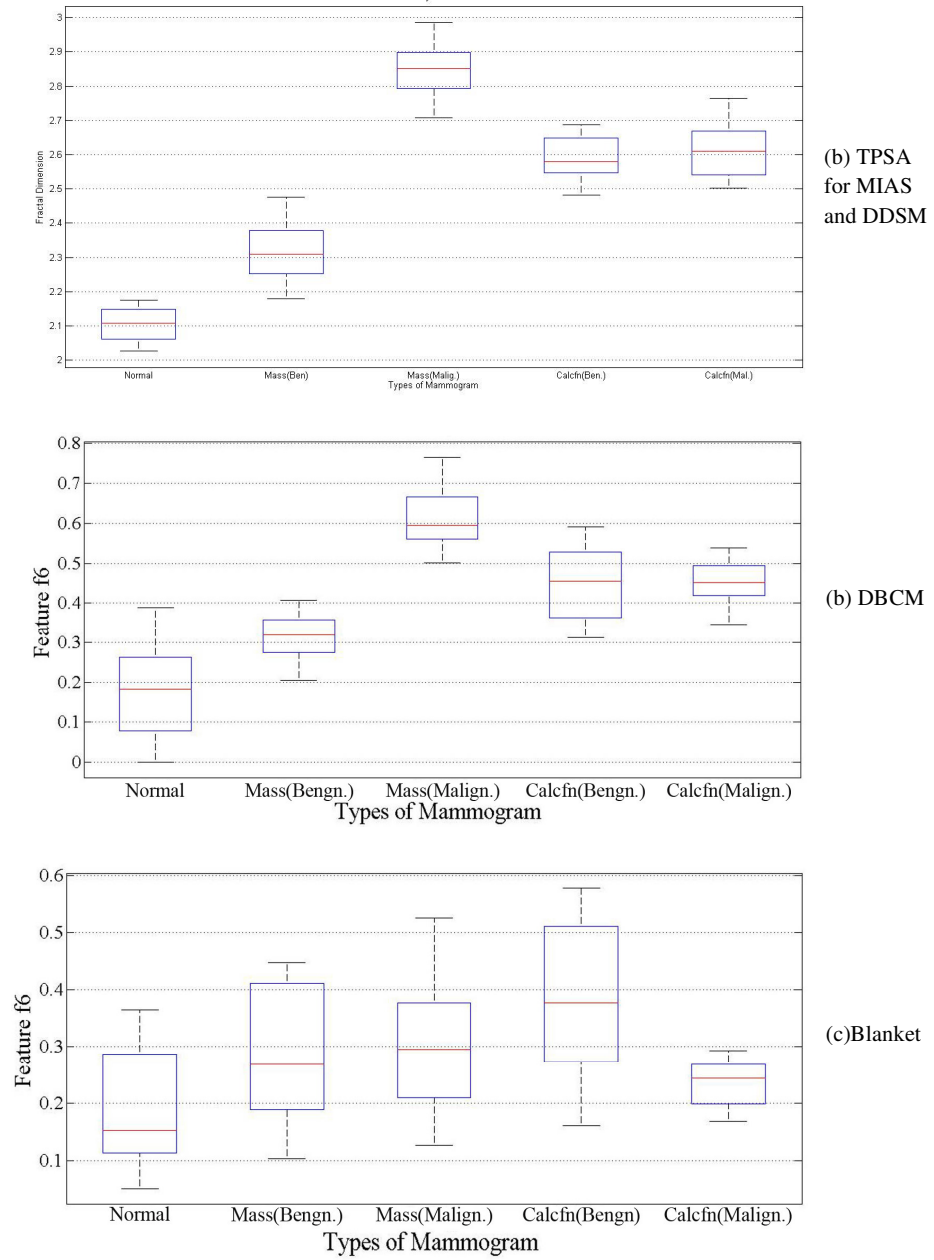


Fig. 4.18 Box Plot of the feature f_6 obtained using TPSA, DBCM and Blanket methods for MIAS and DDSM together

Table 4.25 Mammograms correctly classified using Fractal feature f_6 computed using TPSA, DBC and Blanket methods

Database	No. of Mammograms Classified Correctly by Fractal feature f_6 Computed using			
	TPSA	DBC		Blanket
		Normal	Mass (Malig.)	Normal
MIAS (287 Nos)	All mammograms	66	23	87
DDSM (512 Nos)	All except 11 benign microcalcifications	96	58	104
Overall Classification accuracy%	98.623	30.413		23.905

In the blanket method only the normal mammograms were classified accurately. 87 from MIAS and 104 from the DDSM database were categorized correctly effectively by this feature, giving an overall accuracy of 23.905%. Table 4.25 illustrates the classification accuracies obtained this feature using different methods.

From the above evaluations, it can be concluded that fractal feature f_6 obtained by the TPSA method is the best feature for classifying the different classes of mammograms with a total classification accuracy of 98.623%. But the range of f_6 between benign and malignant microcalcifications is less and so, other features may be added for better estimation, if needed.

4.9.4 Performance Evaluation of the Features

A good feature must be sufficiently discriminating. It was seen that the classification accuracy is the highest with the feature f_6 computed using the TPSA method. These performances were evaluated using the freely available software MedCalc® Version 11.3.3.0. The discriminative powers of these features were found out by performing the t-test on the features and finding the p value. i.e. if the p-values are smaller than 0.05 the features have strong discriminative power. Also the mutual information criterion between different features was also computed to find the best feature. The mutual information criterion was the highest for f_6 , computed using TPSA method. It

was 3.493 and its p-value was 0.0001. Another method to decide the number of features needed for classification is to plot the misclassification error which is the ratio of the number of misclassified observations divided by the number of observations on the test set as a function of the number of features. The plot of the misclassification error is shown in fig.4.19.

The misclassification error (MCE) is plotted along the y axis while the 21 features (FD using TPSA, DBCM, Blanket, $f_1 - f_6$ (TPSA), $f_1 - f_6$ (DBCM) and $f_1 - f_6$ (Blanket) methods) are plotted along the x axis . Only two features FD using TPSA and f_6 using TPSA had the minimum classification error of 9.96% and 1.025% respectively. This substantiates the best feature selected.

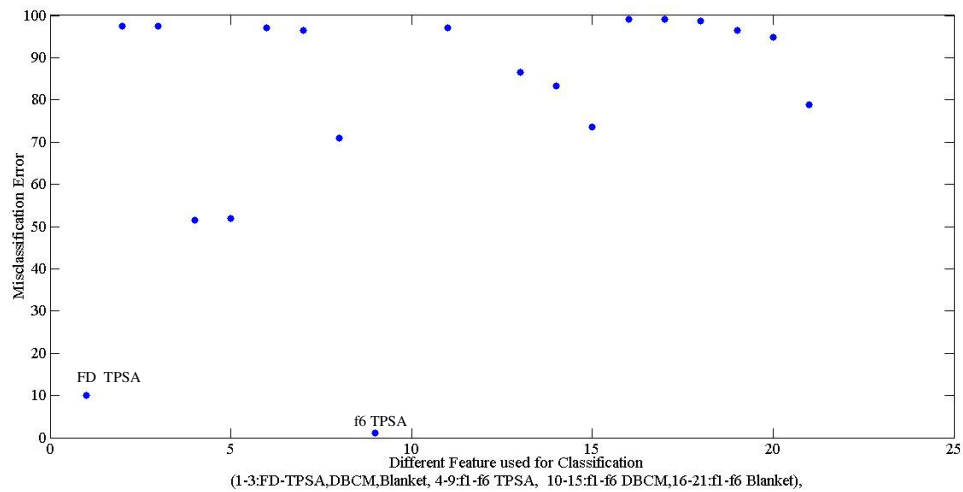


Fig.4.19. Plot of Misclassification Error for different features

4.9.4.1 Statistical analysis

In order to evaluate and compare the discriminative power of these fractal features in distinguishing abnormality from normal breast parenchyma, Receiver Operating Characteristic (ROC) analysis (Mtz 2008), with the parameters defined in section 4.7, was conducted. The ability of a test to discriminate diseased cases from normal cases is based on receiver operating characteristic (ROC) analysis. The ROC curve

represents a plot describing the classifier's true positive detection rate versus its false positive rate.

The false positive (FP) rate is the probability of incorrectly classifying a nontarget object (normal tissue region) as a target object (tumor region). On the other hand, the true positive (TP) detection rate is the probability of correctly classifying a target object as being indeed a target object. Both the TP and FP rates are specified in the interval from 0.0 to 1.0, inclusive. In medical imaging the TP rate is commonly referred to as sensitivity, and $(1.0 - \text{FP rate})$ is called specificity. The schematic outcome of a particular test in two populations, one population with a disease, the other population without the disease, is summarized in Table 4.26

Table 4.26 Definition of the parameters for evaluating the detection accuracy

Test Result	Disease		Sum
	Present	Absent	
Positive	True Positive(TP)	False Positive (FP)	TP+FP
Negative	False Negative(FN)	True Negative(TN)	FN+TN
Sum	TP+FN	FP+TN	

Sensitivity is the probability that a test result will be positive when the disease is present (true positive rate, expressed as a percentage): $\text{Sensitivity} = \text{TP} / (\text{FN} + \text{TP})$. Specificity is the probability that a test result will be negative when the disease is not present (true negative rate, expressed as a percentage). $\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$.

The area under the curve (AUC) is an accepted modality of comparing classifier performance. A perfect classifier has a TP rate of 1.0 and an FP rate of 0.0, resulting in an AUC of 1.0. Random guessing would result in an AUC of 0.5.

Comparison of the ROC curves for the fractal dimension obtained in the three methods is shown in fig.4.20. The Area Under the ROC Curve (AUC) can be used to compare the performance of the features. The ROC curves obtained for the different fractal features are shown in fig.4.21 (a)-(f).

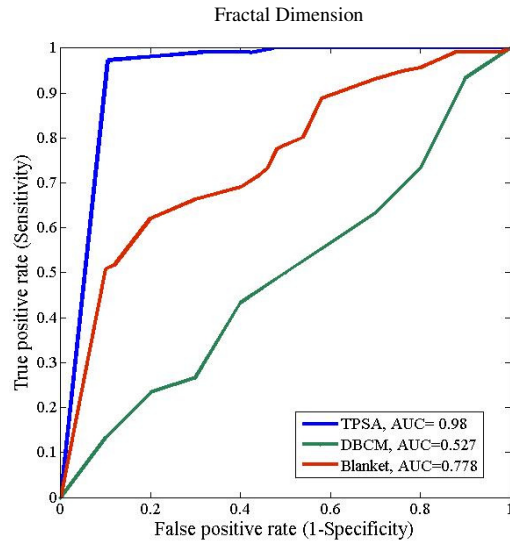
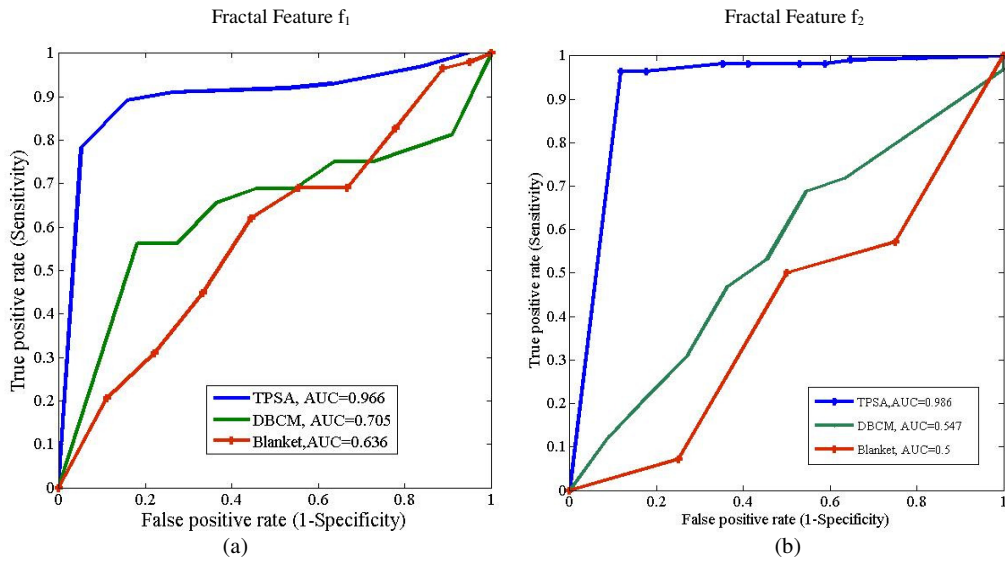


Fig.4.20. ROC curves of Fractal dimension obtained using TPSA, DBCM and Blanket method respectively.



Contd....

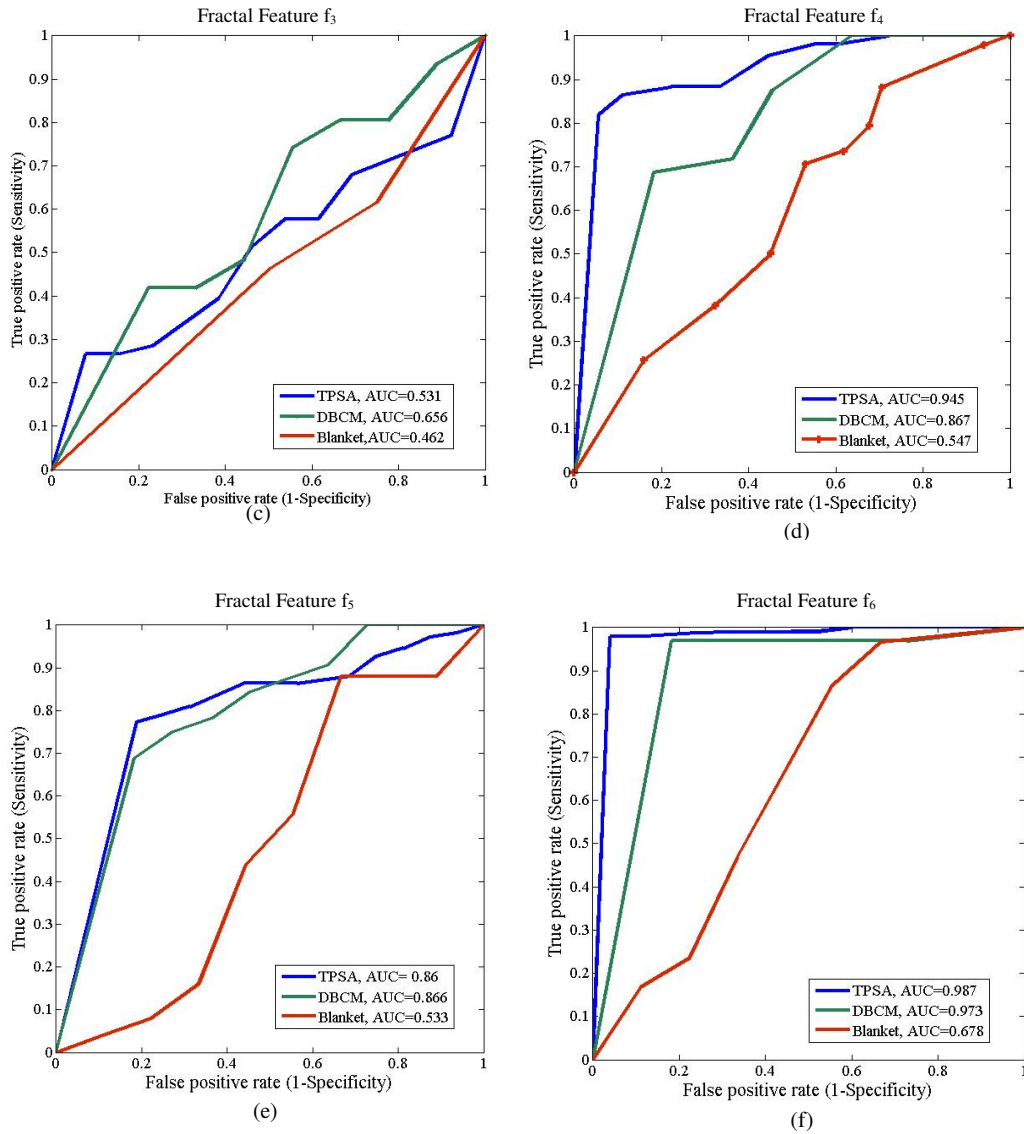


Fig. 4.21(a)-(f) Comparison of ROC curves obtained for the different fractal features

The Area under the ROC curve (AUC), standard error (SE) and 95% confidence interval (Mtz 2008) for the FD and different fractal features computed using TPSA, DBCM and Blanket methods are shown in table 4.27. If independent samples are taken repeatedly from the same population, and a confidence interval calculated for

each sample, then a certain percentage (confidence level) of the intervals will include the unknown population parameter. Confidence intervals are usually calculated so that this percentage is 95%, but it is possible to produce 90%, 99%, 99.9% (or whatever) confidence intervals for the unknown parameter.

The z statistics indicates by how many standard deviations an observation or datum is above or below the mean. The standard error is the standard deviation of the sampling distribution of a statistic. The term may also be used to refer to an estimate of that standard deviation, derived from a particular sample used to compute the estimate.

The values of z statistics obtained between TPSA and DBC methods in computing fractal dimension was 6.014, while that between TPSA and DBC was 4.318. Between DBCM and blanket the z statistics was 0.778. The area under the curve (AUC) is the minimum for the blanket method for all the features. For feature f_1 , the DBC and blanket method had a tendency to misclassify with a lesser amount of Area under the ROC curves (AUC). Due to the differential nature between the minimum and maximum gray levels, DBC method, gave a better classification than blanket method. With feature f_2 , the blanket method had the lowest AUC and cannot classify the different mammograms. It is again observed that TPSA was more suitable than the other two methods. The feature images obtained for estimating f_3 did not contain any information regarding the type of mammogram since there was less variation in the gray level intensity as shown in fig.4.11. The AUC for feature f_3 was found to be very low which is an indication of a poor classifier. Therefore, the classification accuracy of this feature using different FD estimation methods was also poor. The highest AUC was obtained for the DBC method, 0.656 and the lowest was for Blanket method with 0.462. TPSA gave an intermediate AUC of 0.531 with feature f_3 . The 95% confidence intervals for these methods are [0.368, 0.695], [0.464, 0.848] and [0.126, 0.797] for TPSA, DBC and Blanket methods respectively.

The z statistics obtained between the TPSA and DBCM, TPSA and blanket and DBCM and Blanket methods with feature f_4 were 1.345, 3.411, -2.5 respectively. With feature f_5 , the AUC was less for TPSA method. 0.86, 0.866 and 0.533 were the

values obtained with TPSA, DBCM and blanket methods. For feature f_5 , 95% confidence interval of 0.785 to 0.935 was obtained with TPSA and 0.728 to 0.951 was obtained with DBCM. The blanket method had this interval in the range [0.313, 0.754].

The new fractal feature f_6 obtained using the TPSA method gave the highest classification accuracy of 98.623 (Table 4.23). The TPSA method had an AUC of 0.987 with a confidence interval of the mean to be [0.935, 0.999]. For DBCM, these values were 0.929 to 0.99 respectively and AUC was 0.973. AUC of 0.678 with a 95% CI [0.490 - 0.865] was obtained for the blanket method. The z statistics between TPSA and DBCM, TPSA and Blanket, and DBCM and blanket methods were obtained to be 0.148, 2.421 and 2.351 respectively.

Table 4.27

Statistical Analysis of the different fractal dimension and fractal features estimated using TPSA, DBC and Blanket methods

Method	Parameters	FD	Fractal features					
			f_1	f_2	f_3	f_4	f_5	f_6
TPSA	AUC	0.98	0.966	0.986	0.531	0.945	0.86	0.987
	SE	0.007	0.023	0.009	0.084	0.02	0.038	0.010
	95% CI	0.957-0.99	0.921-1.00	0.969-1.00	0.368-.695	0.906-0.984	0.785-0.935	0.995-0.999
DBCM	AUC	0.527	0.705	0.547	0.656	0.867	0.866	0.973
	SE	0.106	0.085	0.999	0.979	0.055	0.055	0.022
	95% CI	0.32-0.734	0.538-0.871	0.351-0.743	0.464-0.848	0.758-0.975	0.728-0.951	0.929-0.99
Blanket	AUC	0.778	0.636	0.5	0.462	0.547	0.533	0.678
	SE	0.036	0.101	0.168	0.171	0.115	0.112	0.096
	95% CI	0.708-0.849	0.437-0.835	0.170-0.829	0.126-0.797	0.367-0.717	0.313-0.754	0.490-0.865

AUC-Area Under ROC curve

SE-Standard Error

CI-Confidence Interval

The ROC also indicates that the feature f_6 computed from the TPSA method was the best compared to the other two methods. Also, f_6 obtained using the blanket method gave a better classification accuracy than the other features computed using this method.

4.9.4.2 Comparison of the performance of the fractal features with conventional features

The classification accuracy of the above introduced fractal features in mammogram classification were compared with the conventional standard intensity, texture, shape features etc discussed in section 4.6. Table 4.28 shows the classification accuracy and the AUC obtained for the conventional features. It was found that these features are having lesser accuracy in classifying mammograms than the fractal feature f_6 .

Table 4.28

Average values of the different conventional features obtained for different classes of mammograms, its classification accuracy and Area Under ROC Curve using same mammograms (287 from MIAS and 501 from DDSM database)

Sl. No.	Features	Normal	Masses		Microcalcifications		Classification Accuracy%	AUC
			Benign	Malignant	Benign	Malignant		
1	Mean	175.48	152.001	164.618	185.724	160.017	43	0.593
2	Skewness	0.122	0.102	-0.037	-0.298	0.238	63	0.586
3	Kurtosis	4.445	3.411	2.664	6.389	5.073	58	0.64
4	Contrast	0.058	0.080	0.083	0.102	0.0589	72	0.68
5	Correlation	0.746	0.800	0.817	0.738	0.873	78	0.7616
6	Cluster Prominence	0.840	1.001	3.088	1.928	1.924	75	0.705
7	Cluster Shade	0.015	0.139	-0.108	-0.166	-0.055	45	0.513
8	Dissimilarity	0.058	0.080	0.082	0.102	0.059	48	0.68
9	Energy	0.699	0.504	0.508	0.554	0.520	81	0.773
10	Entropy	0.582	0.881	0.995	0.925	0.896	83	0.823

11	Homogeneity	0.971	0.87	0.959	0.949	0.971	52	0.687
12	Maximum probability	0.808	0.611	0.671	0.682	0.6628	85	0.815
13	Variance	34.438	29.490	31.086	38.195	28.9032	62	0.6
14	Sum average	11.665	10.842	11.094	12.305	10.7348	49	0.587
15	Sum variance	125.9835	101.141	$\frac{105.57}{7}$	133.908	98.7008	46	0.657
16	Sum entropy	0.5435	0.827	0.943	0.856	0.8600	81	0.84
17	Difference variance	0.0579	0.080	0.083	0.102	0.0588	63	0.68
18	Difference entropy	0.204	0.278	0.284	0.318	0.221	57	0.678
19	Information measure of correlation	0.486	0.695	0.731	0.631	0.774	82	0.863
20	Inverse difference moment normalized	0.999	0.999	0.999	0.998	0.999	38	0.42
21	f_1 (TPSA)	0.1217	0.5087	0.658	0.5578	0.6201	48.686	0.966
22	f_2 (TPSA)	0.1201	0.4962	0.6465	0.5585	0.6197	46.433	0.986
23	f_3 (TPSA)	0.2532	0.2785	0.2894	0.125	0.1324	1.5018	0.531
24	f_4 (DBC)	0.2583	0.332	0.4667	0.450	0.4049	16.27	0.945
25	f_5 (TPSA)	0.0893	0.2087	0.4993	0.177	0.2681	36.54	0.86
26	f_6 (TPSA)	0.0934	0.2841	0.8269	0.4497	0.5861	98.62	0.987

It may be noted that out of the 20 conventional features, maximum classification accuracy and AUC were obtained for the features maximum probability (85 , 0.815), entropy (83, 0.823) and information measure of correlation (82, 0.863), Also, minimum classification accuracy were obtained for inverse difference moment normalized (38,0.42), mean (43, 0.593) and cluster shade (45 , 0.513). The table also illustrates the results obtained with the new fractal features which gave good classification accuracy. f_1 (TPSA) and f_2 (TPSA) gave a classification accuracy and AUC of (48.686, 0.966) and (46.433, 0.986) respectively. Compared to these feature

values, the new fractal feature f_6 introduced in this thesis had a maximum classification accuracy of 98.623% with an AUC of 0.987. This is a good improvement over all other features in mammogram classification.

Based on the results obtained using feature f_6 , a rule based algorithm for the classification of mammogram was formulated as shown in the flowchart given in fig.4.22. Since 11 benign microcalcifications had feature f_6 values overlapping with malignant microcalcifications and the benign microcalcifications, had feature f_6 values between 0.401 and 0.5, this range of value are shown in the flow chart.

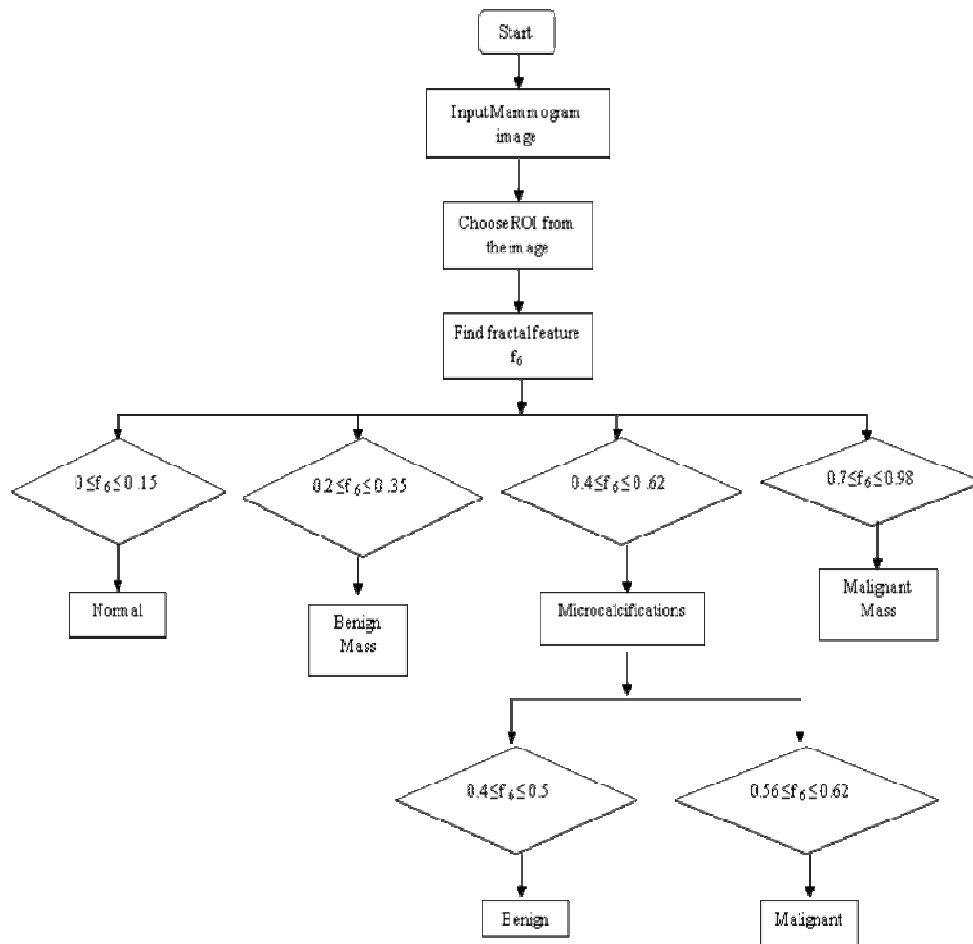


Fig.4.22 Flow Chart for the Classification of Mammograms using fractal feature f_6

4.10 Chapter Summary

In this chapter, a novel fractal feature derived from the TPSA method was developed for classifying mammograms, into normal, benign and malignant masses, and benign and malignant microcalcifications. Different fractal features mentioned by Chaudhuri and Sarkar (Chd 1995) which were used in image segmentation, were used for comparison of classification accuracy. Here, these fractal features are based on the fractal dimension calculated using three methods, viz. Triangular Prism Surface Area (TPSA) method, Differential Box Counting (DBC) method and blanket method. Of the three methods, fractal feature f_6 derived from the TPSA method gave the highest of 98.925% classification accuracy with Malignant masses giving the highest f_6 value in the range 0.7-0.986 and the least feature values of 0.001 - 0.156 were obtained for the normal mammograms as expected. 0.209 - 0.35 is the range of f_6 values obtained for the benign masses while the range of benign and malignant microcalcifications is obtained are 0.401 - 0.5 and 0.56 - 0.62 respectively. Only for benign and malignant microcalcification, there was not sufficient separation between the classes.

The classification performances of these features were evaluated using Receiver Operating Characteristics (ROC) analysis. It substantiates the obtained results. It shows that f_6 has the highest Area Under ROC curve (AUC) of 0.987 and was able to discriminate between different classes of mammograms efficiently. Thus, fractal feature f_6 computed using TPSA method is found to be an effective and promising way to distinguish different types of digital mammograms.

Regular screening of mammograms has been made compulsory in many European countries for women above 40 years of age. Thus the number of mammograms, the radiologists should evaluate will be very large. Also, majority of the mammograms will be normal ones. As a result, radiologists are struggling hard to differentiate between the normal mammograms and the ones with potential breast cancer. In this scenario, radiologists are looking for a more accurate and efficient diagnostic method for detecting cancer cells at an early stage. This study

demonstrates the performance of the new technique developed for the efficient classification of mammograms.

Chapter 5

Identification of Microcalcifications in Mammograms using New Fast Fractal Modeling Approaches

Breast tissue patterns possess local self similarity which is the fundamental property of fractals. Therefore these breast patterns can be easily modeled by fractal method. According to the Collage theorem, modeled image is visually close to the original image. Hence when this modeled mammogram image is subtracted from the original mammogram the presence of microcalcifications, which are early indications of breast cancer can be identified. But, the conventional method of fractal modeling involves a large amount of computation time. This chapter presents four novel fast fractal image modeling approaches for modeling mammograms based on mean variance, entropy, mass center and shade- non shade methods. These methods reduce the time taken to model the mammogram without compromising the accuracy of microcalcification detection. This assists radiologists in identifying microcalcifications in mammograms.

5.1 Microcalcifications in mammograms

Detecting breast cancer at an early stage can help in saving a patient's life. Mammography is long been recognized as the modality for breast cancer screening (Pis 1998). A mammogram can aid in the early detection and diagnosis of breast diseases in women (Rad 2011). The early detection of breast cancer is essential since therapeutic actions are more likely to be successful in the early stages of the disease.

After the classification of the mammograms into different categories, like normal ones, mammograms with benign and malignant masses and mammograms with malignant and benign microcalcifications, as discussed in chapter 4, this research concentrated in identifying the presence of microcalcifications in mammograms.

The analysis of mammographic images is a complex and cumbersome task which requires highly specialized radiologists. Micro calcifications are considered as important early indications of breast cancer. The high correlation between presence of microcalcifications and the presence of breast cancer indicate that the accurate detection of microcalcifications will improve the efficiency of mammogram as a diagnostic procedure. Microcalcifications are very difficult to analyze because of three main reasons:

- a. The size of microcalcifications is very small. They appear as tiny objects which are described as granular, linear or irregular (Li 1997). The size of microcalcifications typically ranges from 0.05 to 1mm. Also, they are embedded in the breast parenchymal textures making them very difficult to detect.
- b. In the mammogram, some parts of the background, such as dense tissue may be brighter than the microcalcifications in the fatty part of the breast.
- c. The intensity and size of the microcalcifications are close to the inhomogeneous background.

Some of the microcalcifications are so faint that they cannot be identified by an experienced radiologist. The subtle and complex nature of the radiographic findings associated with breast cancer causes errors in radiological diagnosis. This

can be attributed to human factors such as varying decision criteria, distraction by other image features and simple oversight (Li 1997).

Mammographic screening programs are carried out in many countries and their effectiveness has a great impact on the prognosis. The screening programs produce a large number of mammograms, majority of which are normal ones. So the screening accuracy of the radiologists can be degraded by the huge case load, which is undoubtedly caused by visual fatigue and other psycho-physical mechanisms. The quality of the diagnostic findings can be improved if double readings are carried out (Boc 2004).

To increase diagnostic efficiency, computer assisted schemes based on advanced image processing and pattern recognition techniques can be used to locate and classify possible lesions. This will alert the radiologist to examine these suspicious areas with particular attention. Moreover, these computer-assisted schemes can improve the performance of the automatic computer-aided diagnosis systems, which can serve as a “prereader” to the radiologist and also give the radiologist a “second opinion” in the diagnosis.

5.2 Fractal Modeling of Mammograms

One of the inherent properties of deterministic fractals is that they have exceptionally high visual complexity with very low information content (Man 1982). These mathematical objects have a high degree of redundancy so that they can be recursively made of transformed copies of either themselves or parts of themselves. Hence, they can be described and generated by simple recursive deterministic algorithms. As seen in chapter 3, images can be modeled by deterministic fractal objects which are attractors of sets of two dimensional affine transformations. This means that an image can be constructed by a set of model parameters which require fewer bits to describe than the original image. The mathematical theory of Iterated Function Systems (IFS) along with the Collage theorem constitutes the broad foundations of fractal image modeling.

It is noted that most of the techniques used for microcalcification enhancement in the past research works, not only detected microcalcifications, but also enhanced background structure and noise. It is observed that microcalcifications are visible as small objects which appear to be added to the mammographic background. Some of these objects are bright, some are faint. Microcalcifications can be characterized in different shapes. But compared with breast background tissue, they have less structural relation. On the other hand, the mammographic parenchymal and ductal patterns in mammograms possess structures with high local self-similarity which is the basic property of fractal objects. Therefore to enhance microcalcifications alone, in the present work, the breast background information is modeled using the deterministic fractal coding method (Li 1997). The disease patterns, i.e., microcalcifications, alone will then be enhanced when the background is subtracted from the original mammograms.

Fractal Image Coding finds a lot of applications in image processing. The next section (5.3) gives the literature survey conducted in the area of fractal image coding.

5.3 Literature Survey

Fractal image coding based on the fractal theory of iterated contractive transformations was developed by Barnsley (Bar 1988). This approach relies on the assumption that image redundancy can be efficiently captured and exploited through piecewise self-transformability on a block-wise basis and it approximates an original image by a fractal image, obtained from a finite number of iterations using an image transformation called a fractal code. This approach was later modified by his student Jacquin (Jaq 1992, Jaq 1993) as fractal block Coding.

The fractal image coding method with the iterated function systems has been widely used in literature for image compression (Hrt 1996, Cab 2002, Zao 2005, Zha 2000, Tho 1995, Jia 1995).

The conventional fractal image coding method was modified and was applied in several image compression applications (Mon 1994, Dvn 1996, Sau 1996, Eug

2006). Comparison of fractal image compression with other compression techniques was dealt with by Davis (Dvs 1998) and Jackson and Hannah (Jac 1993). Hamzaoui modified the decoding algorithm and applied for image compression (Ham 1996). A new technique of fractal image compression for fast encoding using IFS with probabilities was introduced by Mitra *et. al.* (Mit 2001).

An approach to the fractal coding of images was presented by Monro (Mon 1993), in which image blocks are represented by least-squares approximation technique. In results presented by Monro and Dudbridge, (Mon 1992a, Mon 1992b) an image is partitioned into non-overlapping fixed blocks and then encoded without searching, by tiling it with reduced copies of itself and by using a least-squares criterion to derive an optimal mapping. But, this method being an approximation degrades more gracefully as block sizes are increased to gain compression. Monro's algorithm was modified by Pi and Ding (PiM 1997) by defining neighborhoods for the values of fractal codes thus improving the compression ratio and Peak Signal to Noise Ratio (PSNR).

Hurtgen (Hur 1993) presented an algorithm for the exact determination of the Eigen values or the spectral radius of the transformation matrix for a wide range of encoding process and therewith a necessary and sufficient criterion for convergence of the reconstruction process. Overlapping domain blocks of size 12×12 were chosen and by starting the new domain after every 4 pixels in the coding method suggested by Beaumont (Bea 1990) gave compression of 0.8 bit per pel, without introducing obvious artifacts to the untrained eye.

A classification based on the contraction characteristics of transformations in fractal image coding was presented by Yisong *et. al* (Yis 2002). They have found that it will effectively solve the contradiction of match parameter selection and PSNR value and generates a much higher rebuilt image quality with the same or less coding time. Different fractal image coding techniques were compared by Polvere and Nappi (Pov 2000) based on the speed and the quality of the image obtained after decoding. The method based on the mass center feature vector gave a faster coding method with less coding time.

Lu (Lu 1996) compensated the redundancies like pixel translation, rotation, brightness changes or a combination of these operations, which helped in achieving better compression ratio. Cochran *et. al.* (Coc 1996) extended fractal image compression to three dimensions and was used for the analysis of CT and MR images of the head.

Fractal image coding finds application in other areas of image processing as well. Ida and Sambonsugi (Ida 1995) applied fractal coding to image segmentation in which the encoding method used the same conventional fractal coding method. Here a discrete dynamical system is formed with the fractal code. Then a relationship between the regions of objects in an image and trapping regions on the dynamical systems were defined. An image can then be segmented by calculating basins on the dynamical systems. Hanqiang *et. al.* (Hqg 1998) used the mean square error distance (MSED) between the range and its corresponding domain block while fractal encoding to extract edges in images. Since this method avoids gradient operation it is insensitive to noise.

Fractals are used in image enlargement by Chung *et. al* (Chu 2003). The affine transformations were used for texture coding by Malassenet (Mal 1993). Bolotov and Tkach (Bol 2007) used Fractal Ultra Wide Band (FUWB) signal to ensure high immunity against different types of noise existing in communication channels which prevents unauthorized access and may be further utilized in FUWB radar. Image enlarging can be achieved with fractal image coding by scaling the initial image to the desirable size at the very beginning and then scaling the size of range blocks and domain blocks with the same scale factor at each iteration. This method was improved by introducing an additional enhancement layer for each scaled range block in the contractive mapping. This layer helps in preserving the edge regions without destroying the smoothness of the flat regions.

GA (Genetic Algorithm) based fractal coding method to find the optimal IFS parameter with minimum reconstruction error was reported by Takezawa *et.al* (Tak 2003) and higher-quality images at the very low bit-rates below about 0.035-0.05bpp were obtained. Thus their method was projected as a powerful image coding

technique for the wireless communication which requires high performance at the low bit-rates.

The inverse problem involved in fractal image coding deals with finding the IFS parameters of a signal that is exactly generated via IFS. Rinaldo and Zakhor (Rin 1994) used wavelet transforms and image moments to solve the inverse problem and then approximation in the least square sense.

Cochran *et. al* (Coc 1996) extended several fractal image compression techniques on volumetric data, and a new 3D edge classification scheme based on principal component analysis was performed.

Vinoy *et al.* (Vin 2003) considered fractal dimension as a mathematical property in the design and optimization approaches of fractal shaped antennas. They relate, multiple resonant frequencies of a fractal element antenna using Koch curves to the fractal dimension of the geometry.

A new monofractal model to capture scaling behavior of the internet traffic by fitting the cumulants of the measured traffic to the cumulants of the process generated by the model was developed by Molnar and Terdik (Mol 2001).

An attempt has been made to measure the fractal dimension of real visual textures as functions of scale, and to compare the results with synthetic (genuinely) fractal surfaces was presented by Dennis and Dessipris (Den 1989). The imaging surfaces generated in computer graphics are been modeled using the scale invariance property by Kube and Pentland (Kub 1988).

Adaptive contrast enhancement using multi scale representations like microcalcifications, masses, and spicules were done by Laine *et. al* (Lai 1994).

5.4 Mathematical Foundations for Fractal Image Modeling

The two basic theorems explained in chapter 3 (sections 3.4.4 and 3.4.5), the Iterated Function Systems (IFS) and Collage theorem forms, the broad foundations for fractal modeling of images. The theoretical explanation begins with definition of the affine

transformation which is already given in chapter 3. Here is an elaboration of those mathematical fundamentals when applied for fractal image modeling.

The heart of fractal image modeling is the finding of affine transformation of an image. An affine transformation $\omega: R^n \rightarrow R^n$ can be written as $\omega(\vec{x}) = A\vec{x} + \vec{b}$, where $A \in R^{n \times n}$ is an $n \times n$ matrix and $b \in R^n$ is an offset vector. Such a transformation will be contractive exactly when its linear part is contractive, and this depends on the metric used to measure distances. If a norm $\|\cdot\|$ is selected in R^n , then $\vec{x} \rightarrow A\vec{x}$ is contractive when

$$\|A\| = \sup_{x \in R^n} \frac{\|A\vec{x}\|}{\|\vec{x}\|} < 1 \quad (5.1)$$

Let (X, d) denote a metric space of digital images, where X is a set, $d: X \times X \rightarrow R$ is a given metric. Given a complete metric space (X, d) , a metric space $(H(X), h)$ is defined, where $H(X)$ is the space of compact subsets of X and the distance $h: H(X) \times H(X) \rightarrow R$ between two sets A and B is the Hausdorff distance, which is characterized in terms of the metric d [Bar 1988].

Now let $I \in H(X)$ be an original image to be modeled. The chore is to construct a contractive image affine transformation ω defined from $H(X)$ to itself, for which I is an approximated fixed point, which is called an attractor. Thus, find transformations, $\omega: H(X) \rightarrow H(X)$ satisfying the requirement

$$\exists s < 1, \quad \forall I_1, I_2 \in H(X), h(\omega(I_1), \omega(I_2)) \leq sh(I_1, I_2) \quad (5.2)$$

such that

$$h(I, \omega(I)) < \delta \quad (5.3)$$

where δ is a tolerance which can be set to different values according to different applications. The scalar s is called the contractivity of ω . It is shown that ω is a set of

contractive mappings as given in equation (3.14), ω_i , i.e., $\omega = \bigcup_{i=1}^N \omega_i$. According to

the deterministic fractal theory, a set of contractive mappings ω_i is the main part of an Iterated Function System (IFS) as observed in section 3.4.4.

Let $\{X; \omega_i, n=1,2,\dots,N\}$ be an IFS with contractivity factor s . Then the transformation $\omega_i : H(X) \rightarrow H(X)$ defined by $\omega(B) = \bigcup_{n=1}^N \omega_n(B)$ for all $B \in H(X)$, is a contraction mapping on the complete metric space $(H(X), h)$ with contractivity factor s . Its unique fixed point, or attractor, $A \in H(X)$, obeys, equation (3.15), i.e. $A = \omega(A) = \bigcup_{n=1}^N \omega_n(A)$ and is given by $A = \lim_{n \rightarrow \infty} \omega^{0n}(B)$ for any $B \in H(X)$.

ω^{0n} denotes n iterations of the map ω .

The Collage Theorem shows that, once an IFS is found, i.e., ω is known such that $h(I, \omega(I)) < \delta$ is satisfied, then from any given image I_0 and any positive integer n one can get

$$h(I, \omega^{0n}(I_0)) \leq \frac{1}{1-s} h(I, \omega(I)) + s^n h(I, I_0) \quad (5.4)$$

Since $s < 1$, after a number of iterations, the constructed image $I_n = \omega^{0n}(I_0)$ will be visually close to the original image I .

The key point of fractal modeling is to explore the self similarity property of images. In the real world, images are hardly ever self-similar, so it is impossible to obtain a transformation ω for an entire image. But all natural images exhibit local self-similarity. To exploit this local self similarity, the image is divided into n small blocks, and for each block, find a corresponding ω_i . Finally, the transformation

$$\omega = \bigcup_{i=1}^n \omega_i \text{ can be defined for a complete image.}$$

Now, as mammograms are gray level images, the derivation of the affine transformation for digital gray level images are considered. The entire image can be represented in a 3D space, the x and y co ordinates representing the spatial co ordinates and the gray level of the image being represented along the z axis. Let $N_1 = [1, 2, \dots, M]$, $N_2 = [1, 2, \dots, N]$ and $N_3 = [1, 2, \dots, L]$ for an image of size

$M \times N$ with L gray levels respectively. For any digital gray-level image $I(k, l)$, $(k, l, I(k, l)) \in N_1 \times N_2 \times N_3$. The fig.5.1 illustrates the mapping of domain and range in an image. For modeling of images using fractals, first the entire image is divided into non overlapping blocks called range blocks, $R = [R_1, R_2, \dots, R_n]$ and overlapping domain blocks, $D = [D_1, D_2, \dots, D_n]$. Then blocks, R and D will be the subsets of $N_1 \times N_2$ such that $\bigcup_{i=1}^n R_i = N_1 \times N_2$. The size of the domain should be larger than the range, and then only the mapping from the domain to range will be contractive. Usually the domain is taken as twice the size of the range. The mapping ω_i from domain to the range can be defined as:

$$\omega_i(I(k, l)) = s_i I(k, l) + o_i \quad (5.5)$$

where s_i is the scaling factor and o_i is the offset factor. They are constants in each range block R_i . For each R_i , find a $D_i \subset N_1 \times N_2$ and $\omega_i : N_1 \times N_2 \times N_3 \rightarrow N_3$ such that

$$d(I|_{R_i}, \omega_i(I)|_{D_i}) \quad (5.6)$$

is minimized. d is chosen as the mean square root metric. If $I_1, I_2 \in X$ are two digital images, then the mean square root d_{rms} metric is given by

$$\begin{aligned} d_{rms}(I_1, I_2) &= \|I_1 - I_2\|_2 \\ &= \sqrt{\sum_k \sum_l (I_1(k, l) - I_2(k, l))^2} \end{aligned} \quad (5.7)$$

5.5 Algorithm for Fractal Image Modeling

With $N_1 \times N_2$ pixels and N_3 gray levels of a digital image, R and D are collection of subsets of $N_1 \times N_2$ from which R_i and D_i are respectively chosen. The set of R are considered as blocks of 8×8 , 16×16 , 32×32 pixels, which are not overlapping, while the set D consists of overlapping blocks of 16×16 , 32×32 and 64×64 . Thus, the domain blocks have a size twice the block size of the range. This will result in a contraction in the (x, y) plane. This means that every 2×2 pixel area in the domain

corresponds to a single pixel in the corresponding range. Thus, the average of the four domain pixel values are mapped to the area of the range pixel when $\omega_i(I)$ is computed.

For each R_i , search in all D to find a $D_i \in D$ which minimizes equation (5.6). This is equivalent to finding a part of the image that looks most similar to that of R_i , while searching for the best matching domain.

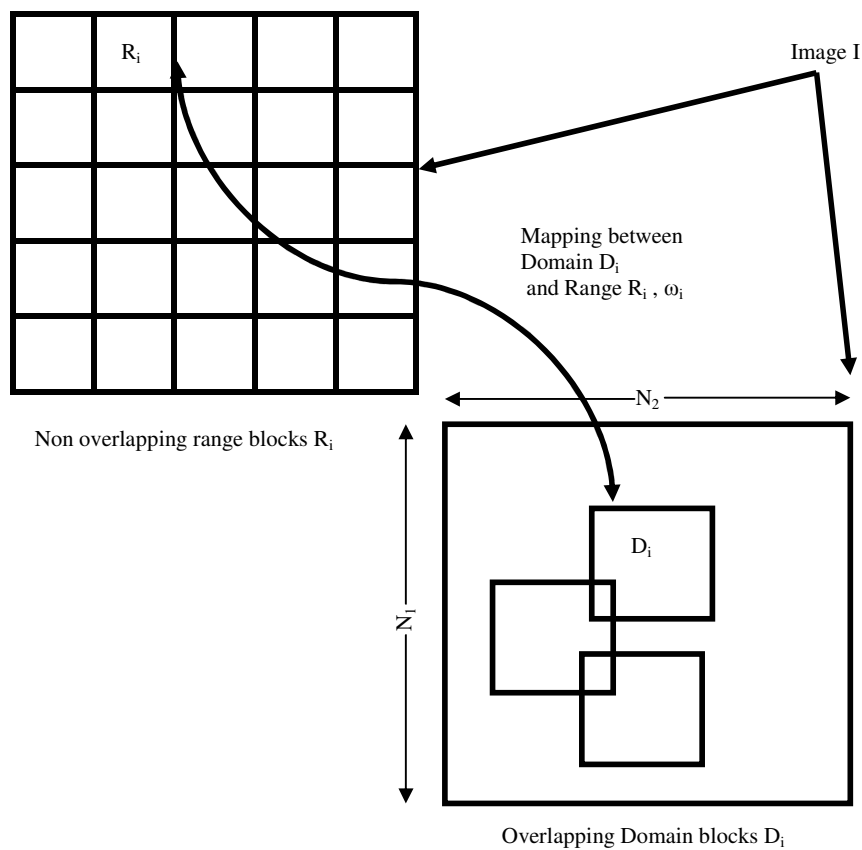


Fig. 5.1 Affine mapping between domain and range

Also, each of the pixel averaged domain blocks can have eight different transformations including flipping and rotation in different orientations, called isometry transformation. They are

- (1) Identity
- (2) Rotation through $+90^0$
- (3) Rotation through $+180^0$
- (4) Rotation through -90^0
- (5) Reflection about mid vertical axis
- (6) Reflection about mid horizontal axis
- (7) Reflection about first diagonal
- (8) Reflection about second diagonal.

There are two steps for the minimization of equation (5.6). At first, optimum values for s_i and o_i should be computed for each ω_i . For each $D_i \in D$, the optimal values of s_i and o_i are computed using the least squares estimation method. From equation (5.6) the unconstrained optimization problem can be constructed as:

$$\min \left\| I(k,l)_{(k,l) \in R_i} - (s_i \bar{I}(k,l)_{(k,l) \in D_i} + o_i) \right\|_2^2 \quad (5.8)$$

where

$$\bar{I}(k,l) = \frac{I(k,l) + I(k+1,l) + I(k,l+1) + I(k+1,l+1)}{4} \quad (5.9)$$

for all $(k,l) \cup (k+1,l) \cup (k,l+1) \cup (k+1,l+1) \in D_i$

Then equation (5.8) can be rewritten as

$$\min e_i = \min \sum_k \sum_l \left(I(k,l) - (s_i \bar{I}(k,l) + o_i) \right)^2 \quad (5.10)$$

By solving $\partial e_i / \partial s_i = 0$ and $\partial e_i / \partial o_i = 0$, the optimal values of s_i and o_i are obtained as follows:

$$\hat{s}_i = \frac{N \sum_k \sum_l I(k,l) \bar{I}(k,l) - \left(\sum_k \sum_l I(k,l) \right) \left(\sum_k \sum_l \bar{I}(k,l) \right)}{N \sum_k \sum_l I^2(k,l) - \left(\sum_k \sum_l I(k,l) \right)^2} \quad (5.11)$$

$$\hat{o}_i = \frac{\sum_k \sum_l \bar{I}(k,l) - \hat{s}_i \sum_k \sum_l I(k,l)}{N}, \quad (5.12)$$

where N is the total number of pixels in R_i . Substituting \hat{s}_i and \hat{o}_i in equation (5.10) obtain the minimum error. Then, a uniform tolerance is set as $\delta_i = \bar{\delta}$ and the best $D_i \in D$ such that $\hat{e}_i < \bar{\delta}$ is selected.

Fractal modeling comprises of two steps: (1) Finding the parameters s_i and o_i from the image is called fractal encoding (2) obtaining the modeled image from these parameters is called decoding.

5.5.1. Algorithm Implementation for Mammogram modeling

In mammograms, microcalcifications appear as a cluster or as single isolated ones on the above described image blocks R_i . In this research, the breast background tissues are modeled using fractals. The intention is to find an area D_i on which the image has a similar structure, but does not contain a similar microcalcification pattern, so that, when the difference between the original image and modeled image is taken, the locations of microcalcifications can be identified. This means that when searching for D_i , the suitable D_i should not cover the region of R_i . Therefore, while searching for the D_i , the constraint $R_i \cap D_i = \phi$ is imposed.

The different steps involved with encoding and decoding in the fractal modeling algorithm are summarized below:

(a) Encoding

Step 1: Image is divided into non overlapping range blocks R_i of size 8×8 . Start with the top most left block and perform a search for the domain which best minimized the equation (5.10) such that the domain D_i satisfies the condition $R_i \cap D_i = \phi$.

Step 2: If the value of equation (5.10) is less than a predetermined tolerance, then the corresponding D_i and ω_i are stored.

Otherwise, if no matching domain is found, the range square is subdivided into four equal squares. Then domain search is performed for each of the four new range blocks and find the most suitable domains. Fig. 5.2 shows the 2 level of quad treeing process. Here, a range block R_i is divided into four blocks R_{i1} , R_{i2} , R_{i3} and R_{i4} . In the next stage R_{i2} is divided into four blocks R_{i21} , R_{i22} , R_{i23} and R_{i24} . This quad treeing process is repeated until the tolerance condition $\hat{e}_i < \bar{\delta}$ is satisfied, or a range square of minimum size is reached. Here, the minimum is chosen as 2×2 pixels. In the present work, the quad tree process was done twice to find a matching domain.

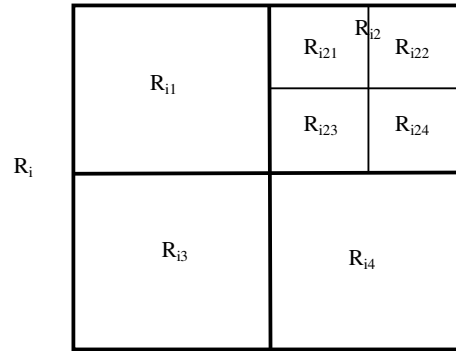


Fig.5.2 Quad tree partitioning of the range blocks

Step 3: The process is carried out until the whole image is modeled. A choice of D_i along with a corresponding s_i and o_i , determines the ω_i on R_i . Once all

ω_i are found, the transformation $\omega = \bigcup_{i=1}^n \omega_i$ can be defined such that

$d(I, \omega(I)) < \delta$ where $\delta = n\bar{\delta}$ and n is the block number of R_i .

Step 4: Finally, based on the Collage Theorem, the modeled image can be easily obtained by performing the iteration on any starting image of the same size according to D_i and ω_i .

(b) *Decoding*

Step1: Start with *any arbitrary image* of the *same size* as that of the original image that was being modeled.

Step 2: For each range, take the stored fractal parameters, s_i and o_i and apply them on the location of their corresponding domain. The pixels of the range location are obtained using equation (5.5). This is repeated for all the range blocks.

Step 3: Go to Step 2. This is done iteratively, till the final modeled image is obtained. Usually convergence was obtained after 12-14 iterations. Even if the number of iterations is increased it does not make any difference in the modeled image. The iteration is stopped when the predetermined tolerance between the original image and modeled image is accomplished.

It is found that only the first four of the isometry transformation given in section 5.5 was sufficient to model and detect the microcalcifications. If all the transformations were used, it was found that there was no improvement in the detection accuracy but the time required for encoding increased.

The flowchart for the above encoding algorithm is shown in fig. 5.3.

5.5.2. Enhancing the Presence of Microcalcifications

Based on the above algorithm, the breast background was modeled. To enhance the presence of microcalcifications the modeled image was subtracted from the original mammogram. But, the direct subtraction does not yield the locations of microcalcification. For that a two step thresholding is done and finally the image is made binary in which the complete background will appear as black and microcalcifications appear as white dots. Let $I(k,l)$ be the original image, and $G(k,l)$ be the modeled image after n iterations. The procedure is summarized as follows.

Step 1: First, the difference between $I(k,l)$ and $G(k,l)$ are found out.

$$I_1(k,l) = I(k,l) - G(k,l), \quad (k,l) \in N_1 \times N_2 \quad (5.13)$$

where $I_1(k,l)$ is the residue image.

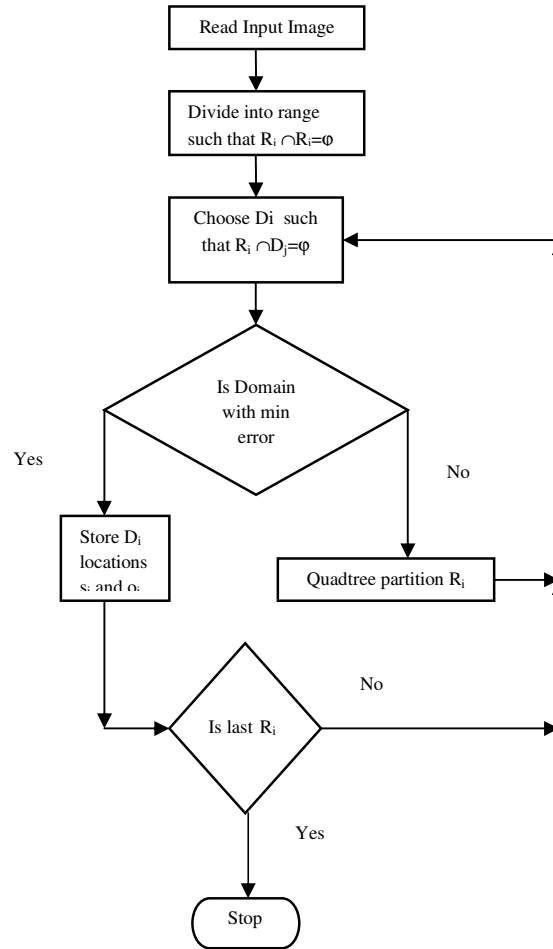


Fig. 5.3 Flowchart for the algorithm for fractal image modeling

Step 2: It is appropriate to ignore the negative value of the difference image $I_1(k, l)$ because negative part of $I_1(k, l)$ does not contain any information about spots including microcalcifications as they are brighter than the background. Therefore

$$I_2(k, l) = \max(0, I_1(k, l)) \quad (k, l) \in N_1 \times N_2 \quad (5.14)$$

where $I_2(k, l)$ is the enhanced image from which the background structures are removed.

Step 3: Image $I_2(k,l)$ contains useful signals and noises. Below a certain threshold T any signal is considered as unreliable. The threshold T is estimated from the image itself as α times the global standard deviation of the noise in the image $I_2(k,l)$. Thus, the value of α is the same for all images, but T depends on each individual image. T can be determined by a two-step estimation process. First the standard deviation of the whole image $I_2(k,l)$ is taken, and the initial threshold T_0 is chosen to be 3.5 times this global standard deviation. Second, only those pixels in which the gray values are below the initial threshold are used to recalculate the standard deviation of the noise. This is one robust method of finding standard deviation of noise (Hub 2009). The final threshold T is determined by adjusting the value of α so that no subtle cases are missed using human judgment and was fixed by trial and error. In this research, α was selected as 0.7. The final enhanced image $I_3(k,l)$ is

$$I_3(k,l) = \begin{cases} I_2(k,l), & I_2(k,l) \geq T \\ 0, & I_2(k,l) < T \end{cases} \quad (5.15)$$

I_3 is the binary image in which microcalcifications appear as bright spots in black back ground.

5.6 Problems encountered during fractal image modeling

The main disadvantage of the fractal modeling method described above is, the enormous time taken during the encoding process. This is because, for every particular range block R_i , all the possible domain blocks available in the image are compared. In the encoding method used in the conventional fractal compression methods, the searching for the next matching domain block starts after leaving a gap of R or $R/2$ pixels from the current domain. But, when this method is used for modeling the mammograms, it was observed that detection accuracy of

microcalcifications was greatly reduced (San 2008b), because the size of smallest microcalcification available in the mammogram database used for the study is 50 μ m.

Therefore to detect even the smallest microcalcification, in this research work, search for the next domain starts from the subsequent adjacent pixel of the current domain itself. This greatly increases the amount of time taken to encode the image. But, this helps in increasing the detection accuracy.

To reduce the enormous time taken for modeling without compromising the detection accuracy four novel fast fractal encoding schemes for modeling the mammograms with better microcalcification detection accuracy was developed.

In the conventional method of fractal image modeling the entire time is utilized for searching a matching domain for a particular range block. While searching for a matching domain, it is unnecessary for the complicated range block, to search all the domain blocks. Instead, if only those domains need to be included in the domain search pool based on certain property in relation to the range. This will reduce the number of isometry transformations required for obtaining a suitable domain. Hence, this helps in reducing the encoding time substantially.

5.7 Fast Fractal Image Modeling

Four methods have been developed to reduce the encoding time during fractal image modeling. In these methods, the entire domain pool is reduced by performing the search based on:

- a. mean and variance
- b. mass center
- c. entropy
- d. shade and non shade blocks

These methods are discussed in the following sections.

5.7.1 Mean and Variance Method

In this method, the property of variance and mean presented by Gi *et.al.* (Gi 2003) is used for distinguishing different image blocks. The variance of block B is given by:

$$Var(B) = \sqrt{\frac{1}{N} \left(\sum_k \sum_l I^2(k,l) \right) - \left(\frac{1}{N} \sum_k \sum_l I(k,l) \right)^2} \quad (5.16)$$

where N is the number of pixels in the range or domain block. The domain block is divided into four sub blocks: top left, top right, bottom left and bottom right as shown in fig.5.4.

The mean value for each sub block is calculated as m_1 , m_2 , m_3 and m_4 respectively. The number of isometry transformations required for finding a matching domain is determined by the class to which these mean values fall into. The transformation algorithm sets four classes as:

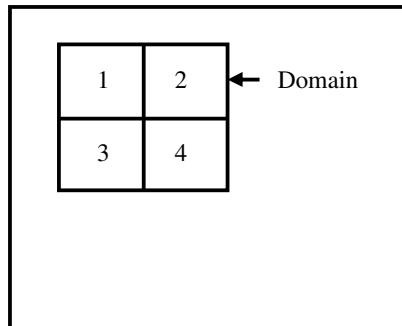


Fig 5.4 Each Domain is divided into four blocks 1, 2, 3, and 4 each of whose mean is computed

Class 1: If $m_1=m_2=m_3=m_4$, it makes only one isometry transformation

Class 2: If $m_1=m_4$ and $m_2=m_3$ it makes two transformations

Class 3: (i) If $m_1= m_2= m_3$ or $m_1= m_2= m_4$ or $m_2= m_3= m_4$ or $m_1= m_3= m_4$ OR

(ii) If $m_1=m_2$ and $m_3=m_4$, OR

(iii) If $m_1=m_3$ and $m_2=m_4$, OR

(iv) If $m_1=m_4$ and $m_2 \neq m_3$, OR

(v) If $m_2=m_3$ and $m_1 \neq m_4$, OR it makes three isometry transformation

Class4: If it doesn't belong to class1, class2, class3, it must make four isometry transformations.

The variances of the range and domain blocks are computed using equation (5.16). For each range block, the domain blocks which satisfy the condition $|\text{var}(R) - \text{var}(\bar{D})| \leq \text{threshold}$, are selected where \bar{D} is the averaged domain block. Then classify those selected domains by their individual mean values. The domain block with minimum mean square error from equation (5.10), is selected from the above domain pool. The location of D_i , values of s_i and o_i form the transformation for the range R_i and they are stored. The process is repeated for all the range blocks in the image.

5.7.2 Entropy Method

Information theory provides the basic tools needed to deal with image representation and manipulation directly and quantitatively. If there exists a set of events

$\{a_1, a_2, \dots, a_n\}$ that occurs with probability $z = \{p_1, p_2, \dots, p_n\}$ where $\sum_{i=1}^n p_i = 1$. Then,

the amount of self information of the event a_i is given by:

$$I(a_i) = \log\left(\frac{1}{p_i}\right) \quad (5.17)$$

The average information content or entropy of each pixel in the image can be computed using the equation:

$$H(z) = -\sum_{i=1}^n p(a_i) \log(p(a_i)) \quad (5.18)$$

As the magnitude of entropy increases, there will be more uncertainty, and hence more information is associated with pixels. It is seen that when domain blocks are analyzed, a large number of domain blocks with high entropy are not used in fractal coding (Has 2005). Therefore, domain pool is constructed with the low entropy

domain, instead of considering all domains. A quality parameter ρ is used to control the number of domains to be included in the domain pool.

5.7.3 Mass center Method

Polvere (Pov 2000) developed mass center features which can be used in fractal image compression techniques. The best matching domain for a range block is the one that has a shape similar to that of the range, up to the isometry transformation. The shape of a block is determined by the distribution of pixel mass within the block, so that similar distribution corresponds to similar blocks. The position of the mass center in a block allows discriminating between blocks that are not similar. The position of the mass center is considered as a feature to search for a matching domain. If two blocks have their mass center close together, then they are similar. For an $N \times N$ block D , the co-ordinates of the mass center are given by,

$$x = \frac{I}{M} \sum_{i=1}^N \sum_{j=1}^N iD(i, j) - \frac{N}{2} \quad (5.19)$$

$$y = \frac{I}{M} \sum_{i=1}^N \sum_{j=1}^N jD(i, j) - \frac{N}{2} \quad (5.20)$$

$$\theta = \tan \frac{y}{x} \quad (5.21)$$

where M is the mass of the block and θ is the mass center expressed in polar co-ordinate system.

If a block D is the matching domain for a range R_i , that is, which minimizes e_i in equation (5.10), then a block $D' = sD + o$, with $s, o \in R$, should also satisfy the above equation. Thus, the two blocks should be characterized by the same features. Applying gray level transformation to a block, causes the block's mass to change by a factor s . The same happens for each pixel in equations (5.19) and (5.20). Thus, it is observed that the co-ordinates x and y of the mass center remains unchanged. When considering the offset part or luminance shifting of the transformation, it causes the mass center of the block to move along the radius connecting the center of the block to the mass center of the original block while the polar angle remains unchanged.

To further characterize the mass distribution of a block D , it is transformed to a new block D_2 according to

$$D_2 = (D - m)^2 \quad (5.22)$$

where m is the average mass of the block D . The features of block D_2 are also calculated using equations (5.19) and (5.20) as above. Thus another feature θ_2 is obtained. With these two features, the feature vector is built as

$$V = (\theta_1, \theta_2) \quad (5.23)$$

While encoding, the feature vector of the range is found as V_r . Compare only those domain blocks whose feature vectors V_d are near to V_r . That is, given a threshold α , only those domains which satisfy the inequality

$$d(V_r, V_d) \leq \alpha \quad (5.24)$$

where d is the distance in the feature space between the two vectors, are compared. A necessary condition for similarity is the suitable distance function given by:

$$d((a_1, a_2), (b_1, b_2)) = \max\{\beta(a_1, b_1), \beta(a_2, b_2)\} \quad (5.25)$$

where

$$\beta(a, b) = \min \{|a - b|, 2\pi - |a - b|\} \quad (5.26)$$

5.7.4 Shade and Nonshade Method

The range blocks R_i are classified into shade and non shade blocks based on their variations in texture (Jaq 1992, Jaq 1993). Shade blocks are those blocks that have no major gradients or texture and the gray scale of pixels change slowly to perception of human eyes. A non shade block has some sudden changes in pixel intensities looking like texture or distinct edges which can be perceived.

Jacquin (Jaq 1992) classified the image into shade, midrange and edge blocks. Mid range blocks are those blocks whose intensity variations falls between shade and edge blocks. In this research, only two classifications were used, i.e shade and non shade, as mammograms are images having low intensity variations and

therefore it is difficult to distinguish between edge and midrange blocks in mammograms. Thus, the classification is limited to shade and non shade blocks.

If the range block is a shade block, no searching is required and only the mean of the pixels is required for decoding. Also, if the domain is a shade block it is not included in the best domain search pool. The non shade range blocks are encoded by the method discussed below.

The transformations that affect the pixels of the transformed domain blocks are considered for this classification. The luminance shift is given by

$$\Delta g = \text{mean}(R_i) - \text{mean}(\overline{D}_i) \quad (5.27)$$

where \overline{D} is the averaged domain block

The contrast scaling factor ν is given by

$$\nu = \min\left(\frac{dr(\text{range})}{dr(\text{domain})}, \nu_{\max}\right), \quad \nu \in [0,1] \quad (5.28)$$

where dr is the dynamic range of the respective blocks, given by the ratio of the maximum pixel value to the minimum pixel value.

$$\text{Dynamic Range of block} = \frac{\text{max value of pixel}}{\text{min value of pixel}} \quad (5.29)$$

Also, the averaged domain blocks can have eight different isometry transformations.

The domain which minimizes the L_2 distortion measure is chosen. The L_2 or root mean square distortion between the image blocks R_i and D_i is defined as the square root of the sum of the squared difference of the pixel values i.e.:

$$d_{L_2}(R_i, \overline{D}_i) = \sqrt{\sum_{k,l} (R_i(k,l) - \overline{D}_i(k,l))^2} \quad (5.30)$$

The fractal coefficients for the range blocks are ν , Δg and isometry value of the corresponding domain along with the domain locations. The fractal code used to represent the entire image is the union of the parameters of all range blocks as follows:

$$\omega = \bigcup_{i=1}^n \omega_i \quad (5.31)$$

5.8 Diagnostic test accuracy for Microcalcification Detection

The diagnostic performance of a test to discriminate between diseased cases from normal case is expressed in terms of four parameters, True Positive (TP), False Positive (FP), False negative (FN) and True Negative (TN). The first letter indicates whether the test result is correct (true) or not (false), and the second letter indicates whether the disease is present (positive) or not (negative); that is, if a normal mammogram is correctly classified as normal, it is called TN, and if incorrectly classified as abnormal, then a FP is counted. Similarly, when a mammogram with microcalcification is correctly classified, then it is a TP; otherwise, if it is incorrectly classified as a normal one, it is a FN. As these terms are used in evaluating the performance of different fractal modeling methods, table 4.26 is reproduced as table 5.1 and tabulates the above discussion.

Table 5.1 Definition of the parameters for evaluating the detection accuracy

Test Result	Disease	
	Present	Absent
Positive	True Positive(TP)	False Positive (FP)
Negative	False Negative(FN)	True Negative(TN)

There are two basic measures of the inherent accuracy of a diagnostic test: sensitivity and specificity. They are equally important, and one is never reported without the other. Sensitivity is the probability of a positive test result (that is, the test indicates the presence of disease) for a patient with the disease. Specificity, on the other hand, is the probability of a negative test result (that is, the test does not indicate the presence of disease) for a patient without the disease. Sensitivity is the probability of a TP among patients with the disease (TPs + FNs). Specificity is the probability of a TN among patients without the disease (TNs + FPs).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (5.32a)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (5.32b)$$

Thus, sensitivity of a diagnostic test is the proportion of patients for whom the outcome is positive that are correctly identified by the test. The specificity is the proportion of patients for whom the outcome is negative that are correctly identified by the test.

The signal to noise ratio (SNR), correlation and the mean square error (MSE) between the original image and the modeled image were used for evaluating the quality of the modeled image. The Peak Signal to Noise Ratio (PSNR) which is used in the analysis is given as:

$$PSNR (dB) = 20 \log_{10} \left(\frac{\text{Variance of the Original Image}}{\text{Variance of the (Original - Modeled) Image}} \right) \quad (5.35)$$

The Mean Squared Error (MSE) gives the pixel by pixel squared difference between the original and modeled images denoted by:

$$MSE = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} (I(i, j) - I_m(i, j))^2 \quad (5.36)$$

where I and I_m are the original and modeled images respectively. Correlation is an indication of the degree of similarity between the two images. Its values are between 0 and 1. For highly correlated images it is close to 1.

5.9 Implementation of Fractal Image Modeling

5.9.1 Database Used

The same mammograms used in chapter 4, obtained from the freely available databases (MIAS) Mammogram Image Analysis Society (Suc 1994) and (DDSM) Digital Database for Screening Mammography (Hea 2001) were used for validating the above discussed methods. The same number of samples from table 4.1 i.e. 166 from MIAS and 180 from DDSM Normal ones and 28 from MIAS and 173 from DDSM mammograms with microcalcifications were used for validating the algorithms developed.

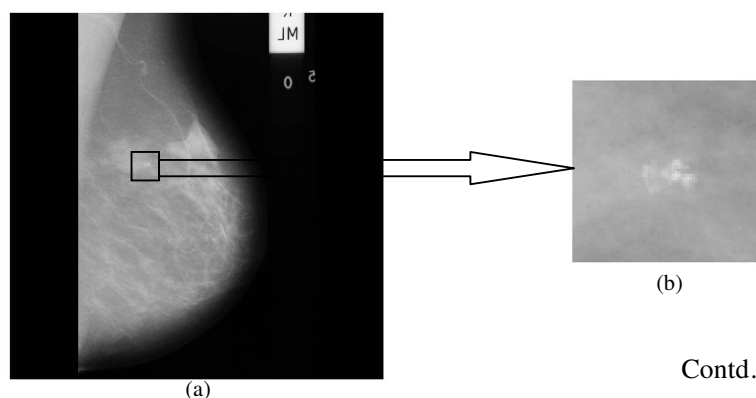
5.9.2 Fixing of different parameters for Fractal Mammogram Modeling

Different Regions of Interest (ROI) containing the microcalcifications were chosen from each mammogram for evaluating and testing the earlier mentioned methods. These ROIs were chosen from the mammogram based on the ground truth available from the databases. For a normal mammogram this ROI was selected arbitrarily. ROIs used for the study were chosen to be 64×64 , 128×128 and 256×256 .

When searching for the most similar domain D_i for a range R_i , only the first four of the isometry transformations, (discussed in section 5.5), of the averaged domain were chosen in all the methods. These four transformations were sufficient enough to detect the presence of microcalcifications. No further improvement in the detection accuracy was obtained when all the eight isometry transformations were utilized.

The size of 1 pixel in the database is $50 \mu\text{m}$. In the fractal method of image compression [Jaq 1992, Jaq 1993], the best matching domain-search was started from the current domain block leaving a gap of R or $R/2$. Small microcalcifications of size one pixel will not be detected if this type of domain search is performed for large R values. Hence in this work, the size of the range blocks used was 2×2 , 4×4 , 8×8 , 16×16 and 32×32 for the above mentioned ROIs. Fig. 5.5 shows the examples of mammograms with different types of microcalcifications.

A region of interest of 64×64 was chosen from each of the mammograms.



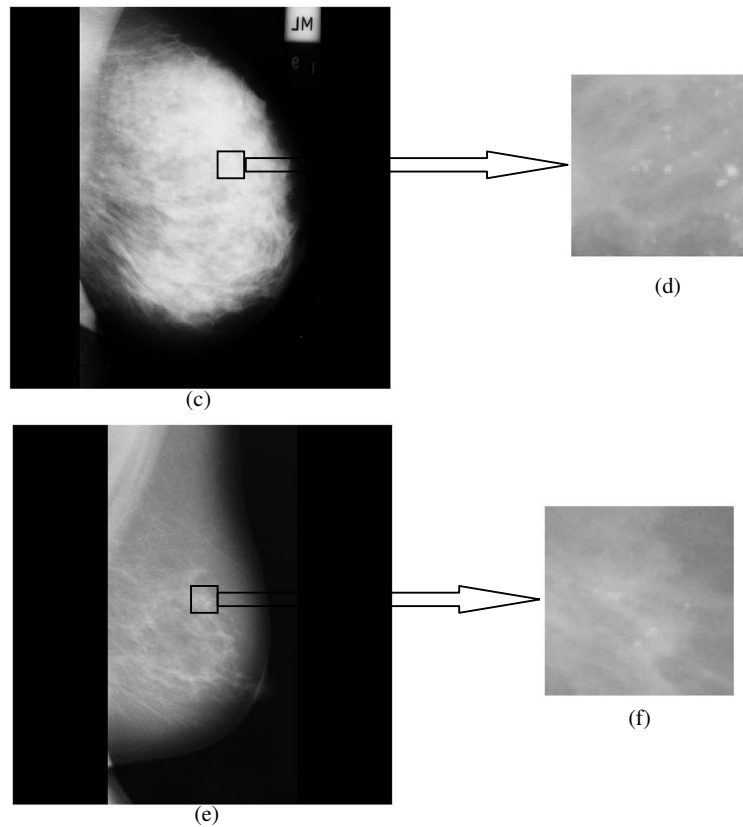


Fig 5.5 Mammograms with microcalcifications (a) Benign Cluster (c) Malignant distributed (e) Malignant Cluster (b) (d) & (f) Region of Interest of 64 x64

Fig.5.5a depicts a typical benign clustered microcalcification in a fatty tissue. Fig.5.5c indicates malignant distributed calcifications. In fig.5.5e, malignant microcalcifications cluster is shown.

In all the fractal modeling methods, the images were normalized before fractal modeling. The value of δ was chosen between 0.01-10. During decoding, different images of the same size, as the original ROI used for encoding, were used as the starting image. Some examples of the starting images used in this study are given in fig.5.6. Fig.5.7 shows the 64x64 ROI chosen from an image, which is used as the starting image in the following discussions.

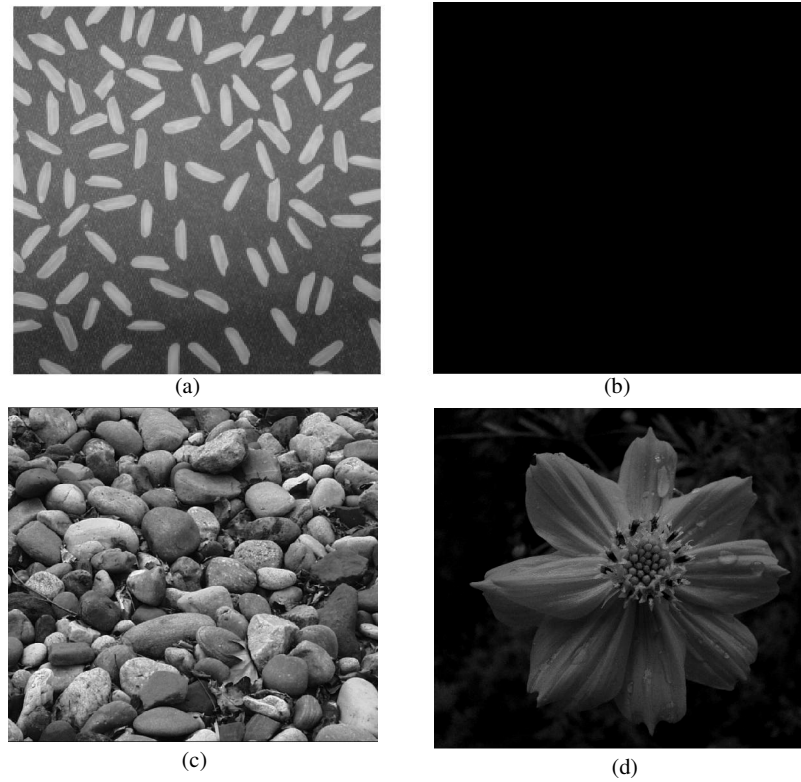


Fig.5.6 Examples of some starting images used during decoding

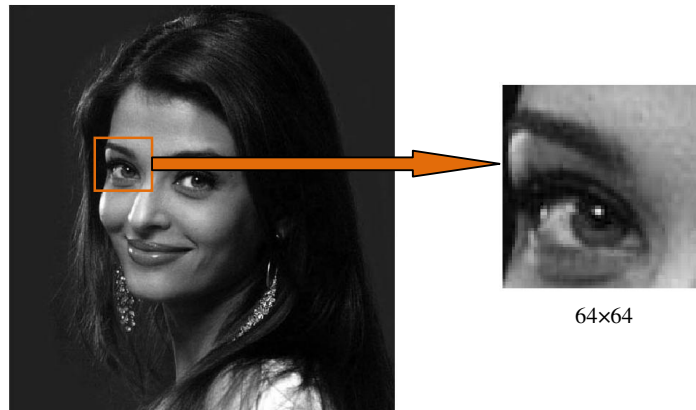


Fig.5.7 Sample 64x64 taken from the original image used as the starting image shown for subsequent modeling figures.

5.10 Results and Discussions

The different fractal modeling methods discussed in sections 5.5 and 5.7 were implemented on mammograms. The algorithm mentioned in section 5.5 was implemented and the different steps involved are:

1. The ROI was chosen to be 64×64 .
2. The image is divided into non overlapping range blocks viz. $2 \times 2, 4 \times 4, 8 \times 8, 16 \times 16$ etc.
3. For each range block the scaling factor s_i and offset factor o_i are found, as per the definitions given in section 5.5.1. These parameters, locations of range and the corresponding domain are also stored.
4. This process is repeated for the entire range blocks.
5. While decoding, any image of the same size as the modeled image can be used as the starting image.
6. The stored parameters are applied iteratively to the arbitrary chosen starting image.
7. The final modeled image is obtained after 10 – 12 iterations in all the methods.

To identify the presence of microcalcifications, the modeled image obtained in step 7 is subtracted from the original image. In the binary difference image, the microcalcifications will appear as white spots in black background. The classification of normal and abnormal was implemented in chapter 4. However, normal mammograms are also used for evaluating the performance of different fractal modeling methods, as these methods are modeling the background of the mammograms. The difference image obtained while subtracting a modeled normal mammogram is completely black, as can be seen in the next sections.

While modeling, the region containing the range is not included in the domain search pool. This is to make sure that the matching domain is not contained in the same region of the range. So microcalcifications which are isolated bright regions will be enhanced when the modeled image is subtracted from the original one. This is applied to all the fractal modeling methods used in this research.

The results obtained after the implementation of these methods are discussed in the next sections.

5.10.1 Implementation of Conventional Fractal Modeling Method

Here, fractal modeling is intended to model the background region in the mammogram and hence normal mammograms are considered for evaluating the quality of the modeled image.

The fig. 5.8 shows the starting image of size 64×64 and result of fractal image decoding after every iteration using the parameters got from normal mammograms.

The ROI chosen was 64×64 and range size chosen was 8×8 . Fig.5.8 (a) shows the original mammogram of size 64×64 to be modeled. Fig 5.8 (b) shows the starting image used for decoding. To this arbitrary chosen image, the parameters obtained from the image to be modeled are applied iteratively. The modeled images obtained after each iteration are shown in figs.5.8 (b)-(o). It is obvious from the figure that convergence is obtained after 10 -12 iterations. The final modeled image in fig. 5.8(o) is visually close to the original image in fig 5.8(a). The difference image is obtained by subtracting the thresholded final modeled image in Fig 5.8 (o) and the original image in fig 5.8 (a) is shown in fig. 5.8 (p). Here, since it is a normal mammogram, the difference image is completely black, indicating that no microcalcifications are present in it.

The fig. 5.9 shows the original ROI of 64×64 chosen from different normal mammograms and the corresponding modeled images obtained after convergence of the iteration. While searching for the matching domain, the next domain is searched from the current domain after leaving a gap of $R/2$.

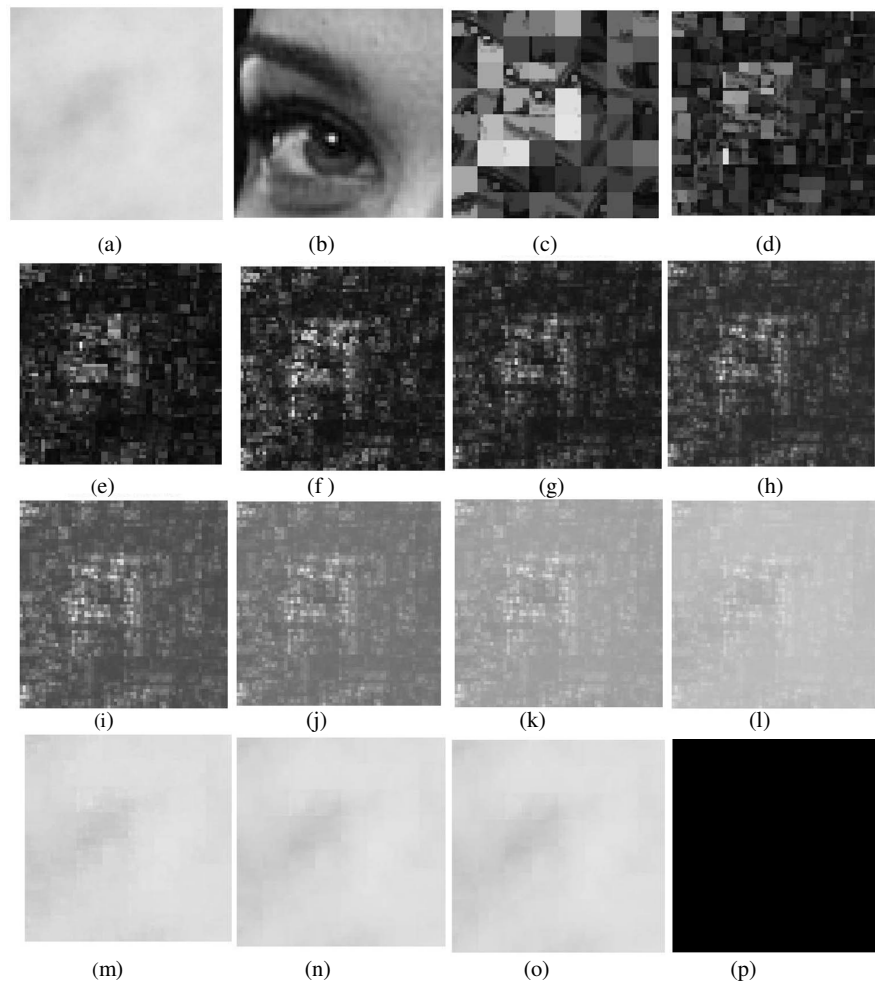


Fig.5.8 Modeling by Conventional Fractal Method (ROI-64×64, range-8×8) (a) Original Image to be modeled (b) Arbitrary Starting Image for decoding (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)

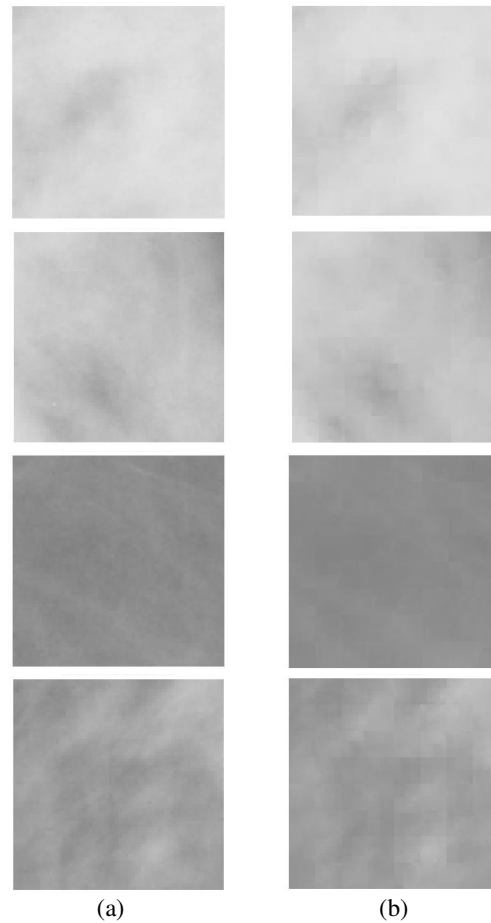


Fig. 5.9 Modeling Normal mammograms by Conventional Fractal Modeling Method
(ROI-64×64, range-8×8)
(a) Original Image (b) Modeled Image

The quality of modeled image is determined by plotting the PSNR, MSE and correlation between the original ROI and the final modeled image obtained after convergence. These parameters are plotted for the conventional fractal modeling method in the fig. 5.10 (a)-(c).

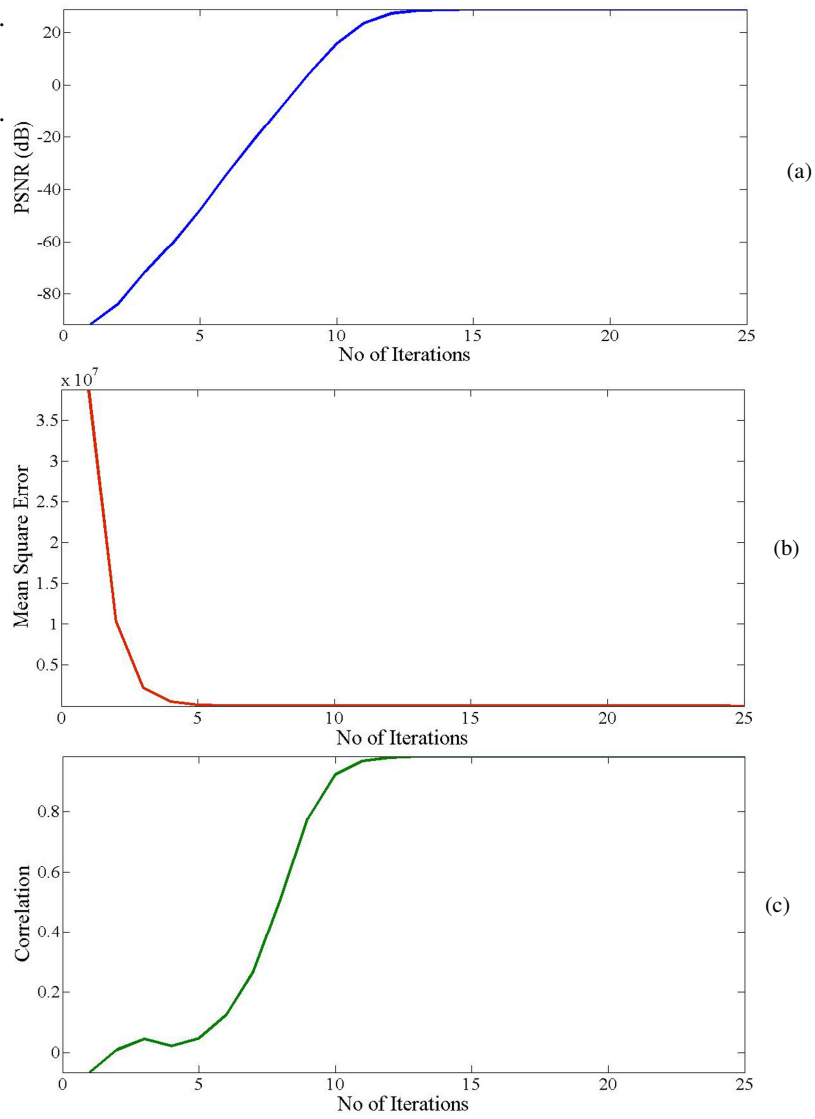


Fig. 5.10 (a) PSNR (dB), (b) Mean Square Error and (c) Correlation respectively between the original and modeled image of normal mammograms for modeling using Conventional Fractal Modeling Method with ROI 64×64 , range size 8×8

In the beginning of the iteration procedure there will be blocking artifacts and therefore the error will be very large. These figures show that after twelve iterations the values of MSE, correlation and PSNR converge to their stable values

Table 5.2 shows average Mean Squared Error (MSE), Correlation and the Peak Signal to Noise Ratio (PSNR) obtained between the original and modeled image while modeling the normal mammograms in the conventional fractal modeling method.

Table 5.2
Evaluating the Modeled image obtained using conventional fractal, modeling method (ROI 64×64, range size 8×8)

Mammogram	MSE	Correlation	PSNR(dB)
Normal	6.783	0.9357	22.76

These results show that the conventional method can model normal mammograms. But, when the modeled image is subtracted from the original there is a chance to miss small microcalcification of the size of a pixel.

The microcalcifications which are detected by subtracting the modeled image from the original image, using conventional fractal modeling method are shown in fig. 5.11. The size of the smallest microcalcification in the data bases is one pixel which is equivalent to 50 μ m. When searching for a matching domain in the domain pool, after checking the current domain, if it does not satisfy equation (5.10), the algorithm searches for the next domain after leaving a gap of R or R/2 in the conventional method. This is suitable for compression applications. But, when the ultimate aim is to detect microcalcifications which are very small in size, of the order of one pixel, domain search of R or R/2 pixels cannot be afforded

Thus, the number of mammograms from which microcalcifications are correctly identified are tabulated in table 5.3 based on the diagnostic efficiency given in table 5.1. As seen in the table 5.3, the microcalcification detection accuracy obtained was only 46.268%. This cannot be tolerated for a good diagnostic test. But, the time taken for the encoding process was very less, around 0.7616minutes.

As the detection accuracy is less than 50%, the conventional fractal modeling method is not used for comparison with the other new methods used for microcalcification detection discussed in the later sections.

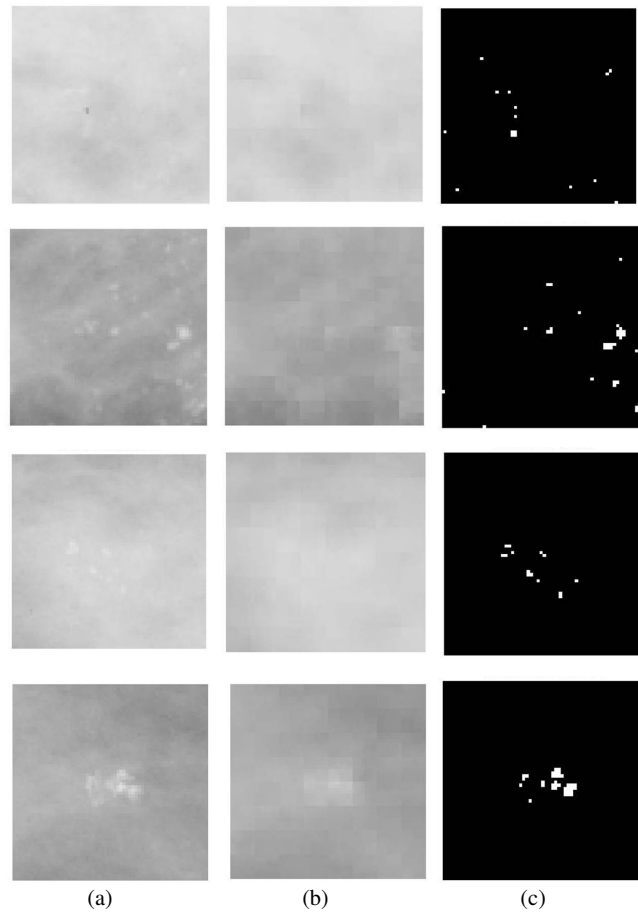


Fig. 5.11 (a) Original Mammogram (b) Modeled Image (c) Detected Microcalcifications using Conventional fractal modeling method (ROI 64×64, range size 8×8)

Table 5.3

Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Conventional Fractal Coding Method

Mammo grams	No. of Samples	Conventional Fractal Coding Method						Time (minutes)	
		TP	FP	FN	TN	Sensi tivity	Specif icity		Avg. No. of Domains searched
Normal	346			41	305		88.15%	430.67	0.7647
With Micro calcifications	201	93	108	-		46.268		466.375	0.7585
Total	547	93	108	41	305	46.268	88.15	448.522	0.7616

5.10.2 Implementation of Modified Conventional Fractal Image Modeling

To improve the microcalcification detection accuracy of the conventional fractal modeling method, it was modified by starting the domain search from the next adjacent pixel of the current domain. This is a novel method of fractal image encoding used in image modeling.

Fig.5.12 shows the original image, starting image used in decoding, the modeled image obtained after different iterations and the difference image obtained after subtracting the modeled image from the original image. Here also the modeling algorithm converged after 12 iterations. During the initial iterations, blocking artifacts are present in the modeled image and it is reduced as the number of iteration increases. The fig. 5.13 shows the modeled images obtained and the corresponding original normal mammograms.

Once a stable modeled image is attained, there are no changes in the MSE, correlation and PSNR between the original and modeled image on further iterations. These values obtained after each iteration while decoding are plotted in fig. 5.14(a)-(c). The average value of the signal to noise ratio, MSE and Correlation between the original normal mammogram and its modeled image is given in table 5.4.

Fig.5.15 shows the microcalcification detection results for the modified conventional fractal modeling method. Fig. 5.15 (a) shows the original image, (b) shows the corresponding modeled image obtained and the (c) gives the microcalcifications which are located when the modeled image is subtracted from the original image.

From the figure it is seen that the microcalcifications which are missed in the conventional fractal modeling method are detected using the new modified method, as more white dots are present in the subtracted image.

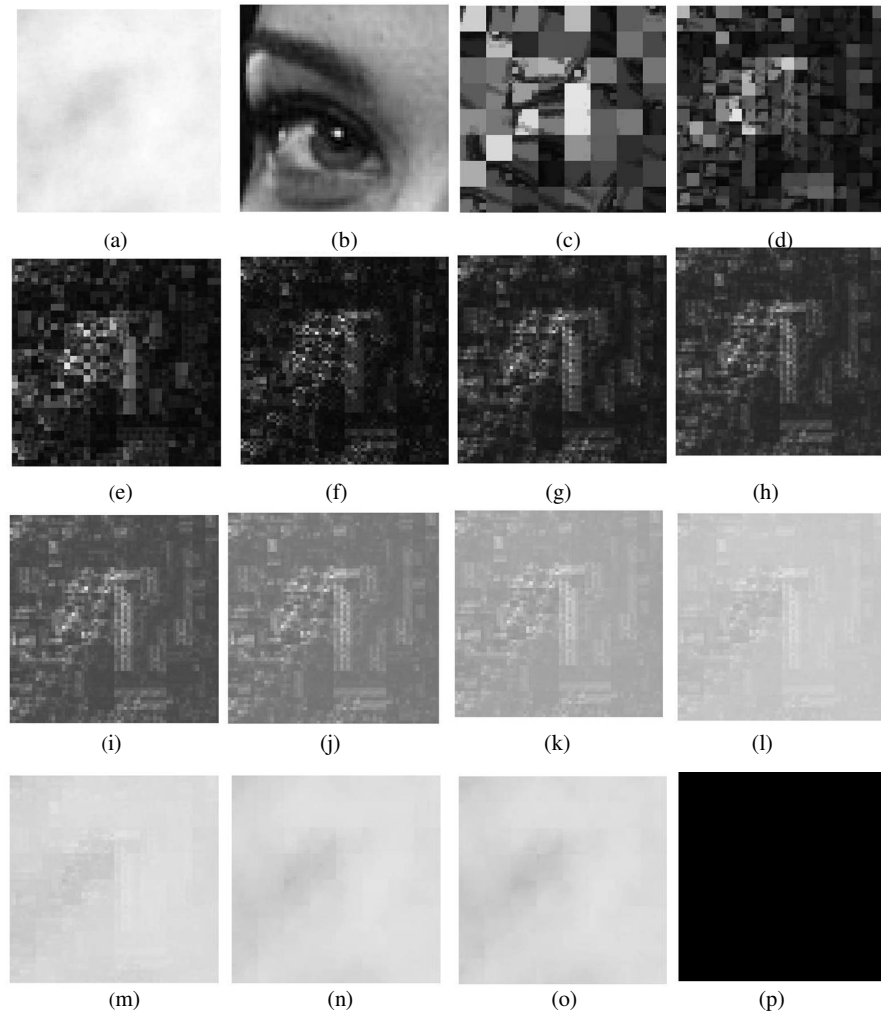


Fig 5.12 Modeling by Modified Conventional Fractal Method (ROI-64×64, range-8×8) (a) Original Image to be modeled (b) Arbitrary chosen Starting Image (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)

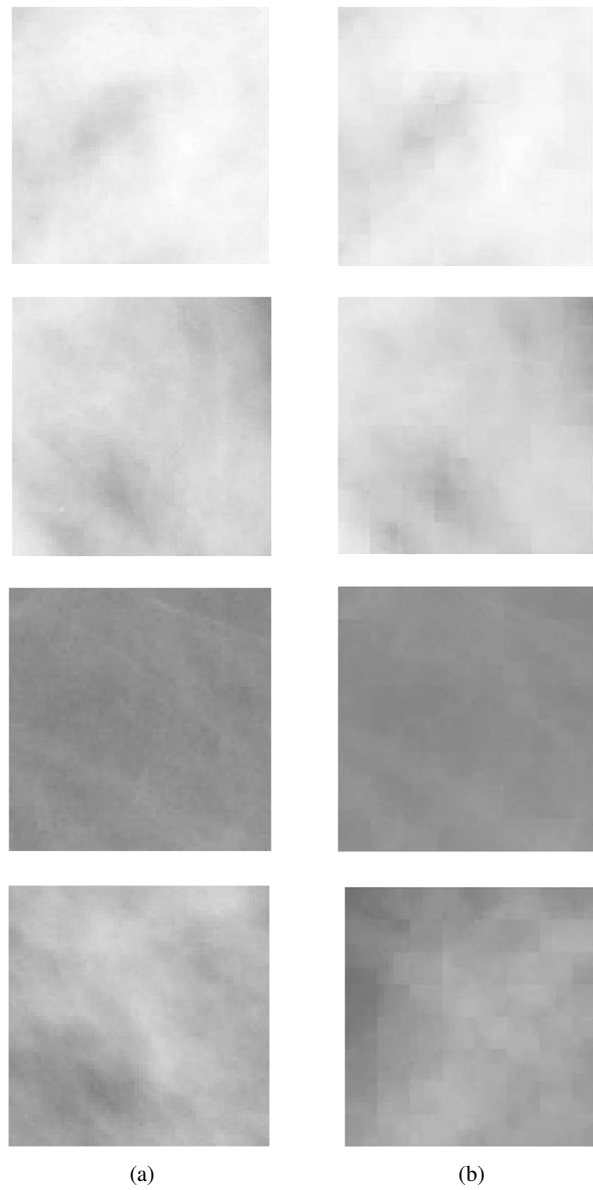


Fig 5.13 (a) Original and (b) Modeled Normal Mammograms respectively by modified Conventional Fractal modeling method (ROI-64×64, range-8×8)

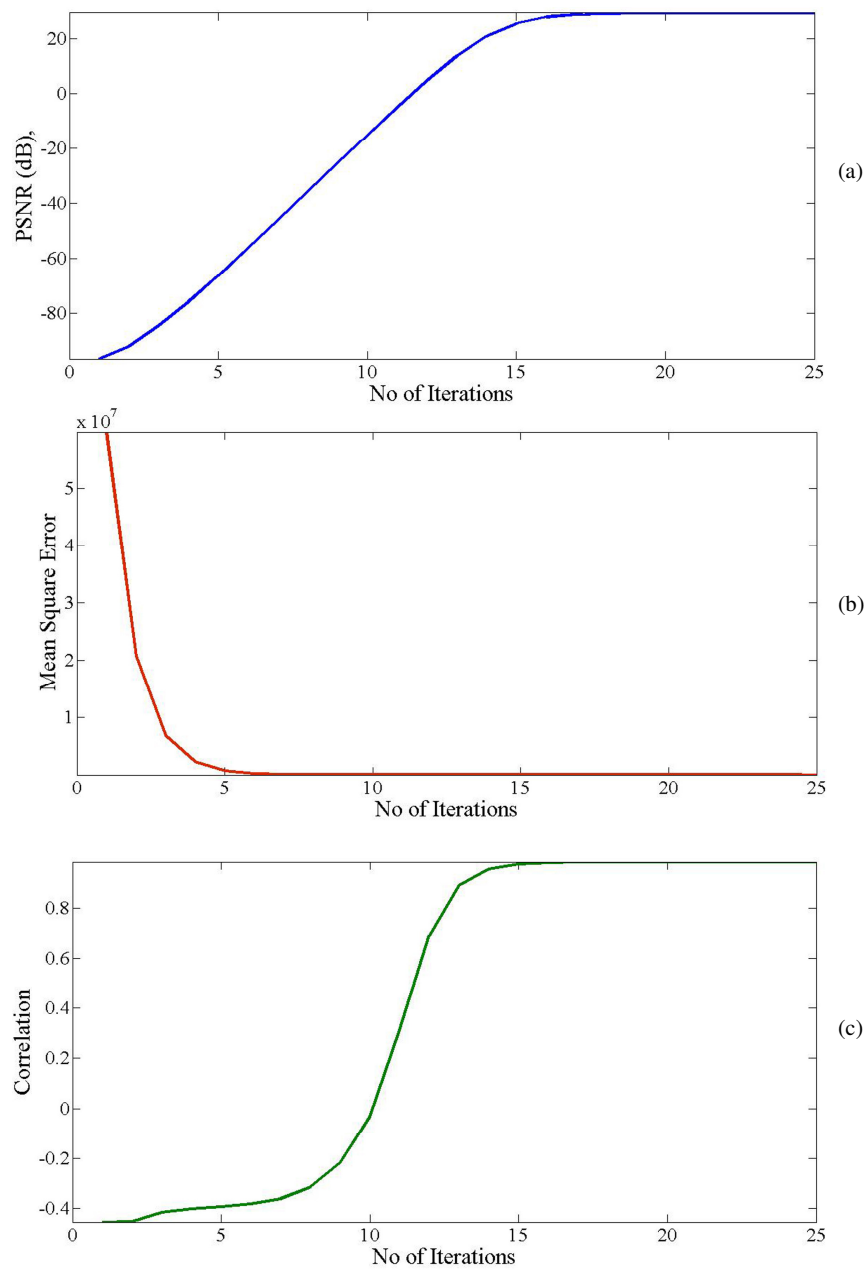


Fig. 5.14 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Modified conventional fractal image coding method with ROI 64×64 , range size 8×8 .

Table 5.4

Evaluating the Modeled Normal Mammogram images obtained using Modified conventional fractal, modeling method (ROI 64×64, range size 8×8,)

Mammogram	MSE	Correlation	PSNR(dB)
Normal	3.154	0.9351	23.783

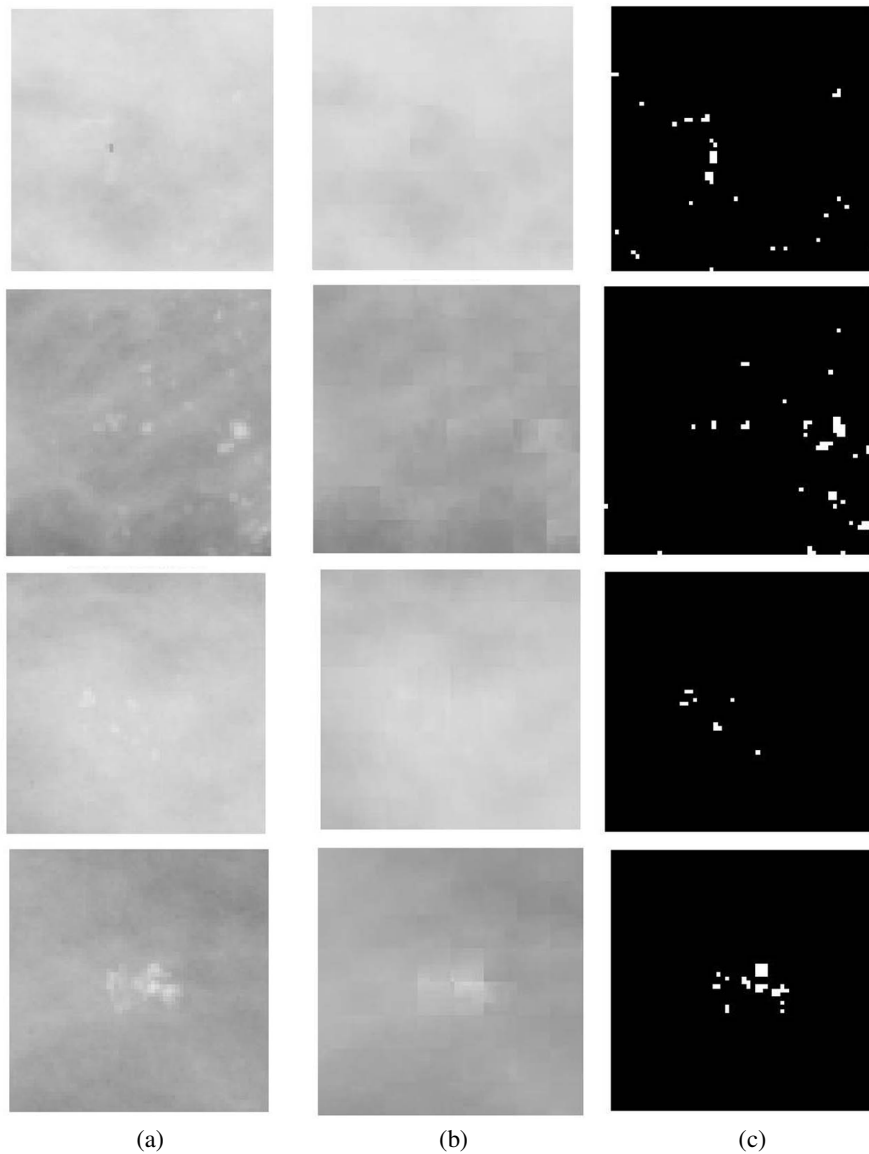


Fig. 5.15 (a) Original Mammogram (b) Modeled Image (c) Detected Microcalcifications using Modified Conventional fractal modeling method (ROI 64×64, range size 8×8).

But, the average number of domains searched per range increased to 1767 from 430.67, for normal mammograms. This tremendously increased the time taken for fractal encoding. The time taken was increased by factor of 34.988 than the conventional method. The detection sensitivity obtained is tabulated in table 5.5.

The number of domain searched per range is in between 1453 and 2228 for and ROI of 64×64, range size of 8×8. For mammograms with microcalcifications, more domains are required to be checked, as discussed in section 5.5.1 microcalcifications appear as clusters or single isolated ones and the region containing the range in the mammogram is not included while searching for the domain. This implies that, to get a matching domain for a particular range block with microcalcifications more number of domains is to be searched. Also, for most of the ranges, quad tree partitioning is necessary to find the suitable matching domain.

Table 5.5

Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for modified conventional fractal coding method, (ROI 64×64, range size 8×8)

Mammo grams	No.of Samples	TP	FP	TN	FN	Sensit ivity (%)	Speci ficity (%)	Avg. No. of Domains searched	Time (minutes)
Normal	346			326	20		94.22	1767	26.756
With Micro calcifications	201	166	35			82.587		2013	25.726
Total	547	166	35	20	326	82.587	94.22	1890	26.241

As the entire possible domain blocks are searched in the modified conventional method, the average number of domain blocks searched per range block is very large, around 3542, for a range size 2×2, 2873 for 4×4 and 1767 for 8×8 respectively. It is found that as the number of domains to be searched per range black increases, the time taken for encoding has become very large when compared to the conventional fractal coding method. But, the microcalcification detection accuracy was significantly increased from 46.268% to 82.587%. The overall detection accuracy was obtained was 89.945% as compared to the 67% of the conventional fractal modeling method. Table 5.6 shows the average number of domains searched

per range and the time taken for the different size of range block. The same detection accuracy was obtained with 2×2, 4×4 and 8×8 range blocks also.

Table 5.6

Time taken and the average number of domains searched for different range sizes for Modified Conventional Fractal modeling method (ROI 64×64)

Size of range	Time Taken (minutes)	No.of Domain searched /range
2×2	315	3542
4×4	93	2873
8×8	26.241	1767

But, when the range size increases above 8×8, blocking artifacts are present in the modeled image and microcalcifications are not effectively identified when the modeled image is subtracted from the original one. Hence, for this method, the range size of 8×8 was chosen to be feasible, with best detection accuracy and less time than other block sizes.

In this method, since the domain search started from the next adjacent pixel onwards, the time required for encoding increased. The disadvantage with this modified conventional method is that the time taken increased. To reduce the time required for encoding without compromising the detection accuracy four new encoding methods were implemented.

5.10.1.3 Implementation of Mean Variance Method

The mean variance method discussed in section 5.7.1 was applied for modeling mammograms. The original normal image, starting image used in decoding, the modeled image obtained after different iterations and the difference image obtained after subtracting the modeled image from the original image are shown in fig.5.16. In this method also, the convergence was obtained after 10 - 12 iterations. Even if the iterations are repeated again, the PSNR, correlation and MSE between the modeled

image and original image remained constant. Fig.5.17 illustrates the original and modeled image of normal mammograms. The variation of MSE, correlation and PSNR with iterations are shown in the fig.5.18 (a) - (c). The average values of these parameters obtained are shown in table 5.7.

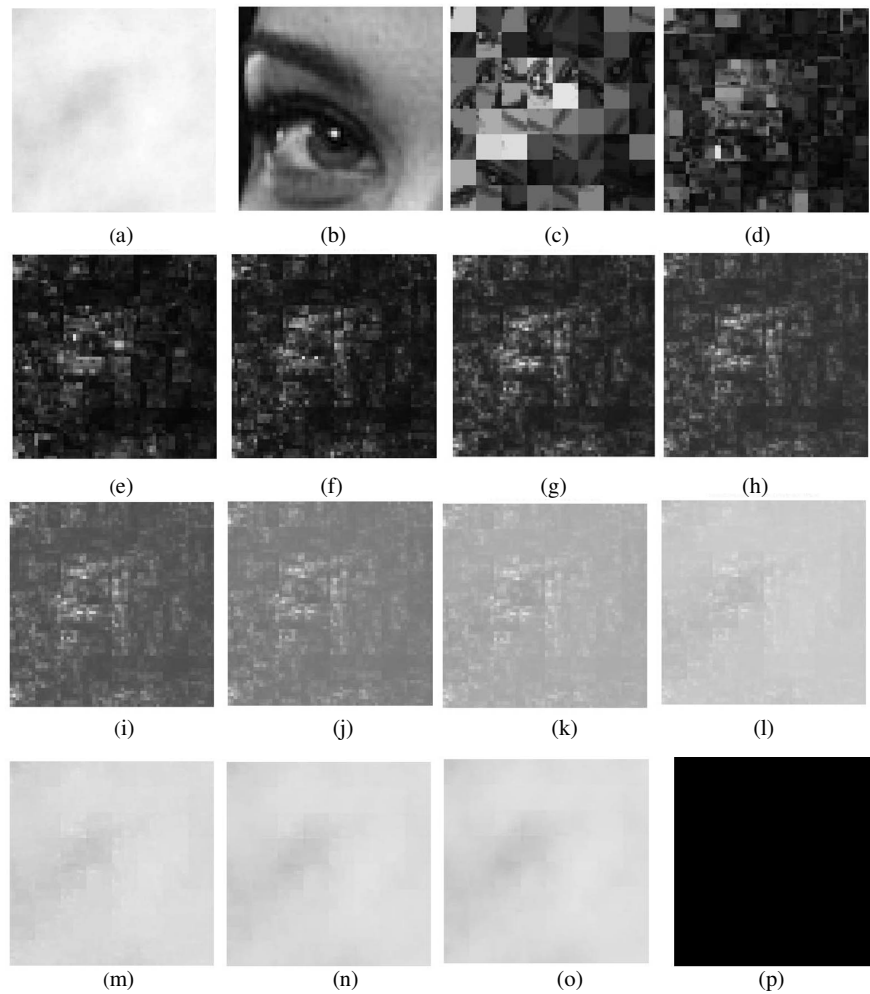


Fig.5.16. Modeling by Mean Variance Method (ROI-64×64, range-8×8) (a) Original Image to be modeled (b) Arbitrary chosen Starting Image (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)

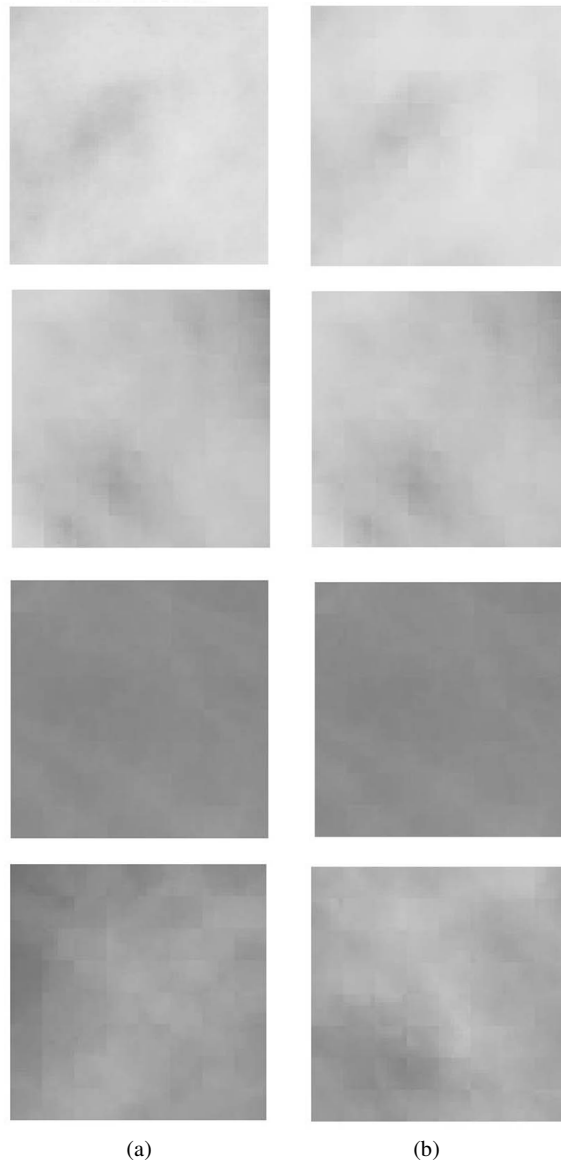


Fig.5.17 Original and Modeled Normal Mammograms by mean variance method
(ROI-64×64, range-8×8)

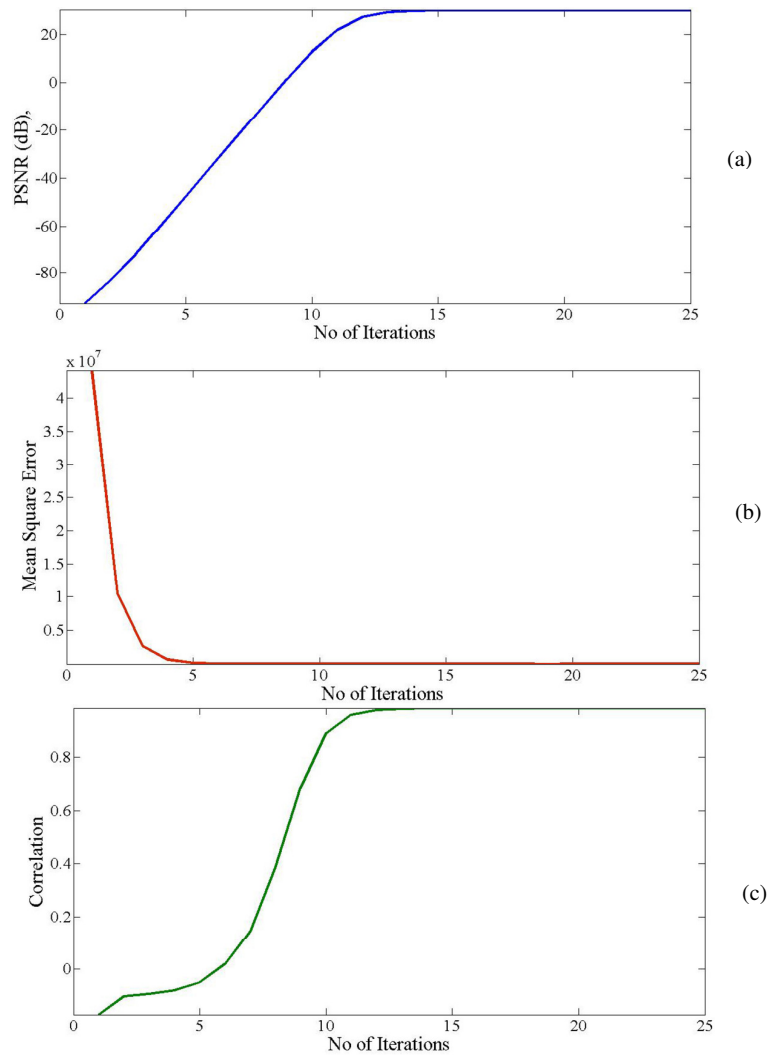


Fig.5.18 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Mean Variance fractal image coding method with ROI 64×64 , range size 8×8 .

Table 5.7 Evaluating the Modeled image obtained using Mean Variance fractal modeling method

Mammogram	MSE	Correlation	PSNR(dB)
Normal	3.756	0.9594	42.5626

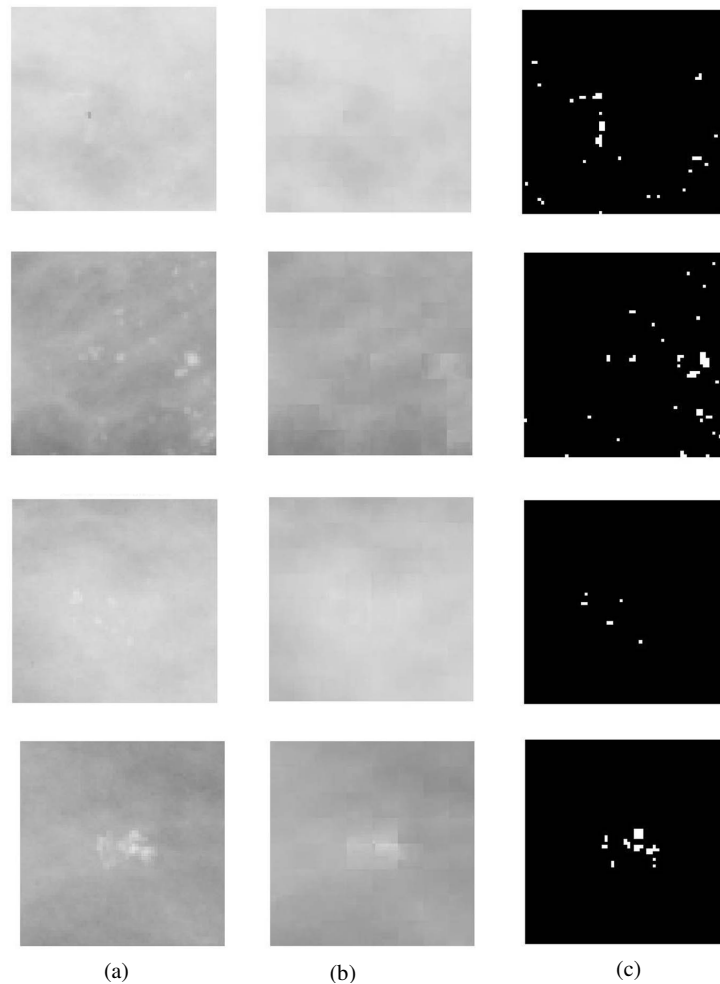


Fig.5.19 (a) Original Mammogram (b) – (c) Modeled Mammogram and Detected Microcalcifications using Mean Variance Method(ROI-64×64, range-8×8)

Fig.5.19 gives the original and modeled image and identified microcalcifications using the mean variance method. For a normal image, if correctly modeled, a complete black image is obtained without any white spots, indicating no microcalcifications were present in it. Microcalcifications of size of a single pixel can be detected by this method. The variance difference was chosen as between 0.001 to 0.8 by trial and error. When this value was above 0.8, blocking artifacts were present in the modeled image and all the microcalcifications could not be detected. The best

detection accuracy was obtained when the range size was 8×8 , as in the modified conventional fractal modeling method. The average number of domains searched per range block for finding a matching domain was obtained as 915.67. This reduced the time taken for modeling by a factor of 3.11. The results obtained with the mean and variance method of classification of domains is given in the table 5.8. Table 5.9 illustrates the time taken and number of domains searched per range block for different range sizes. Same detection accuracy was obtained for all the range sizes.

Table 5.8 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Mean Variance fractal coding method (ROI 64×64 , range size 8×8)

Mammo grams	No. of Samples	Mean Variance Method							
		TP	FP	TN	FN	Sensitivity %	Specificity %	No. of Domains searched /Range	Time (minutes)
Normal	346			326	20		94.219	915.67	8.59
With Micro Calcifns.	201	166	35			82.587		947.46	8.207
Total	547	166	35	326	20	82.587	94.219	931.565	8.399

Table 5.9 Time taken and Average number of domains searched for different range sizes in Mean Variance method (ROI 64×64 , range size 8×8)

Size of range	Time Taken (minutes)	Avg. No. of Domain searched
2x2	22.1801	120.1162
4x4	15.6178	10.65
8x8	8.399	915.6724
16x16	1.0538	306.2857

The time taken for encoding reduced to 8.399 minutes for the range size of 8×8 . But, when the range was increased beyond 16×16 blocking artifacts were present in the modeled image and all the microcalcifications could not be detected.

To further reduce the time and improve the detection accuracy, the following methods were tried.

5.10.1.4 Implementation of Entropy Method

For the entropy based method, the optimum value of the quality parameter ρ was found to be 4. Only those domains which have $\rho > 4$, were included in the domain search pool.

In this method also, the results obtained are explained, by starting with the modeled image of a normal mammogram obtained after each iterations as shown in fig.5.20. It is seen that convergence of the modeled image was achieved after 12 iterations. The difference binary image obtained after the subtraction of the modeled image from the original image is a completely black image indicating the absence of microcalcification in it. The modeled images obtained for different normal mammograms are presented in fig.5.21. The plot of the variation in MSE, correlation and PSNR are shown in fig.5.22. And there were no further changes in these values with iterations once a stable modeled image is attained. The average values of these parameters after convergence is tabulated in table 5.10.

The microcalcifications were identified by subtracting the modeled image after applying the necessary threshold given in section 5.5.2 is shown in fig.5.23.

The diagnostic accuracy obtained with this method of modeling mammograms and the time taken are shown in table 5.11. This reduced the average number of domains searched for each range from 1727 in the modified conventional fractal coding method to 385, reducing by a factor of 4.485. The encoding time is reduced from 26.756 minutes in the modified conventional method to 7.786 minutes in entropy based method.

The results obtained for different range size in ROI of 64×64 is tabulated in table 5.12. The average number of domains searched per block for a range size of 2×2 is 768.403. Because of this, it took 71.806 minutes for encoding.

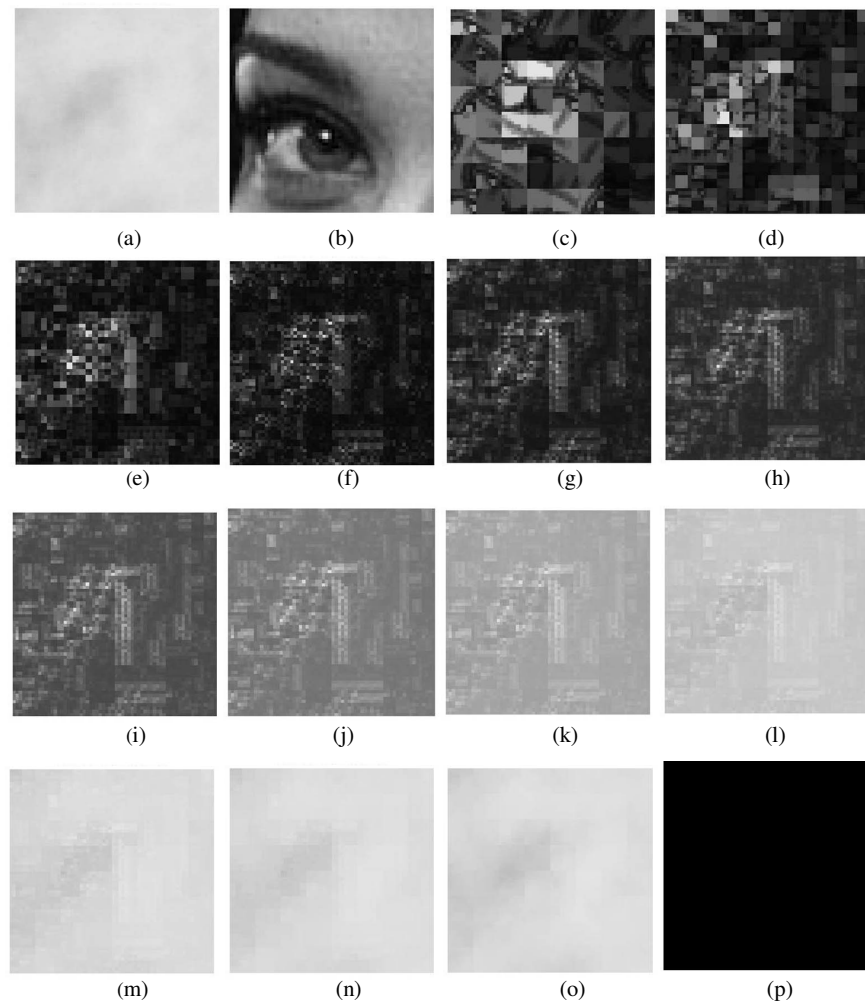


Fig.5.20 Modeling by Entropy Method (ROI-64×64, range-8×8) (a) Original Image to be modeled (b) Arbitrary chosen Starting Image (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)

For a range of 4×4, the average number of domains searched per range block became 412.355. Therefore, the time for encoding was reduced to 13.02 minutes. Compared to the modified conventional method, the average number of domains to be searched per range block is less. But the time taken is more because more number of isometry transformations and quad tree partitioning are required for finding a matching

domain. But as the range size was increased beyond 8×8 , the time required for encoding was reduced. But blocking artifacts were present in the modeled image. But presence of microcalcifications could not be detected correctly when the range size was increased beyond 16×16 .

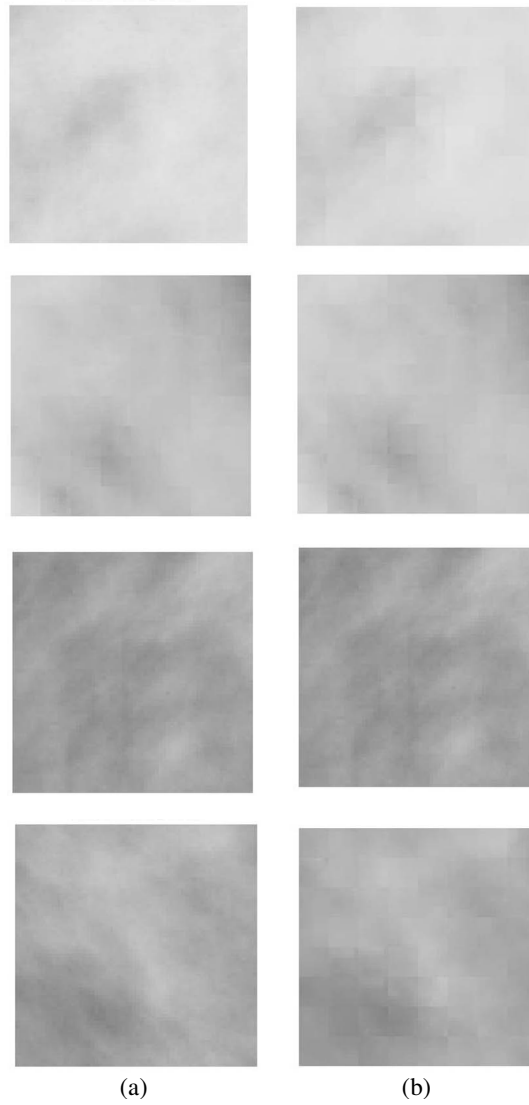


Fig.5.21 (a) Original Normal Mammogram (b) Modeled Mammogram using Entropy Method (ROI 64×64 , range size 8×8)

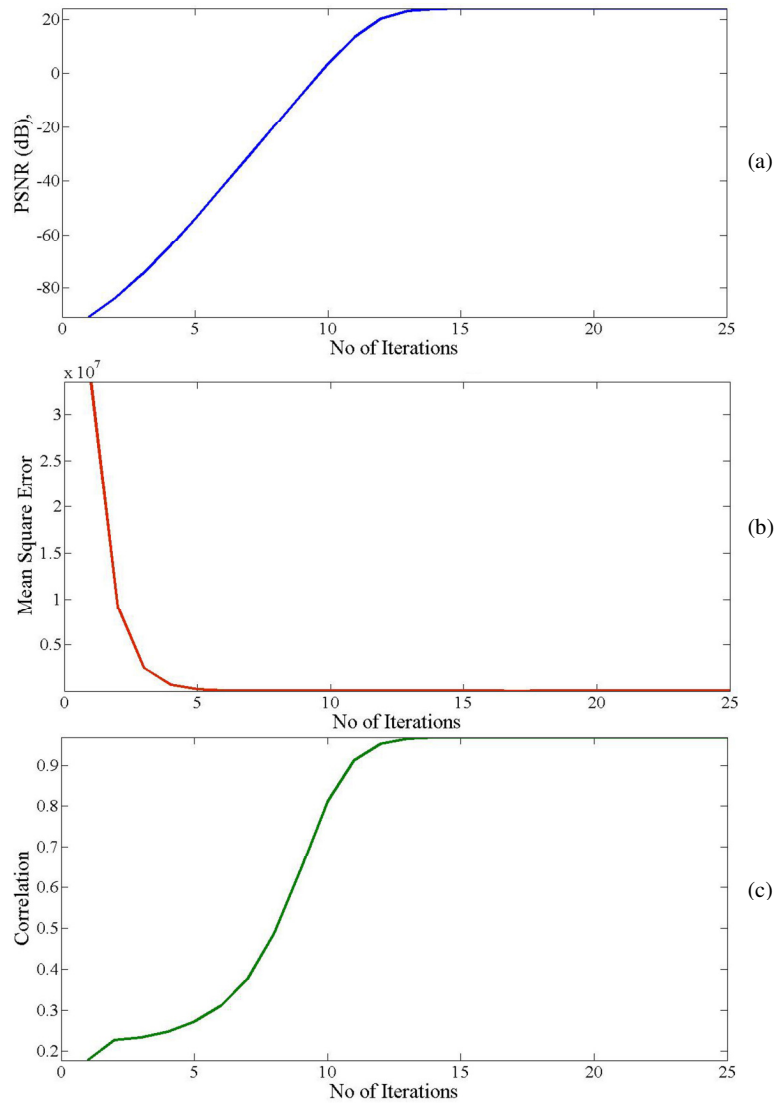


Fig.5.22 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Entropy coding method with ROI 64×64, range size 8×8.

Table 5.10 Evaluation of the Modeled image obtained using Entropy fractal modeling method

Mammogram	MSE	Correlation	PSNR(dB)
Normal	2.6767	9.9752	26.2121

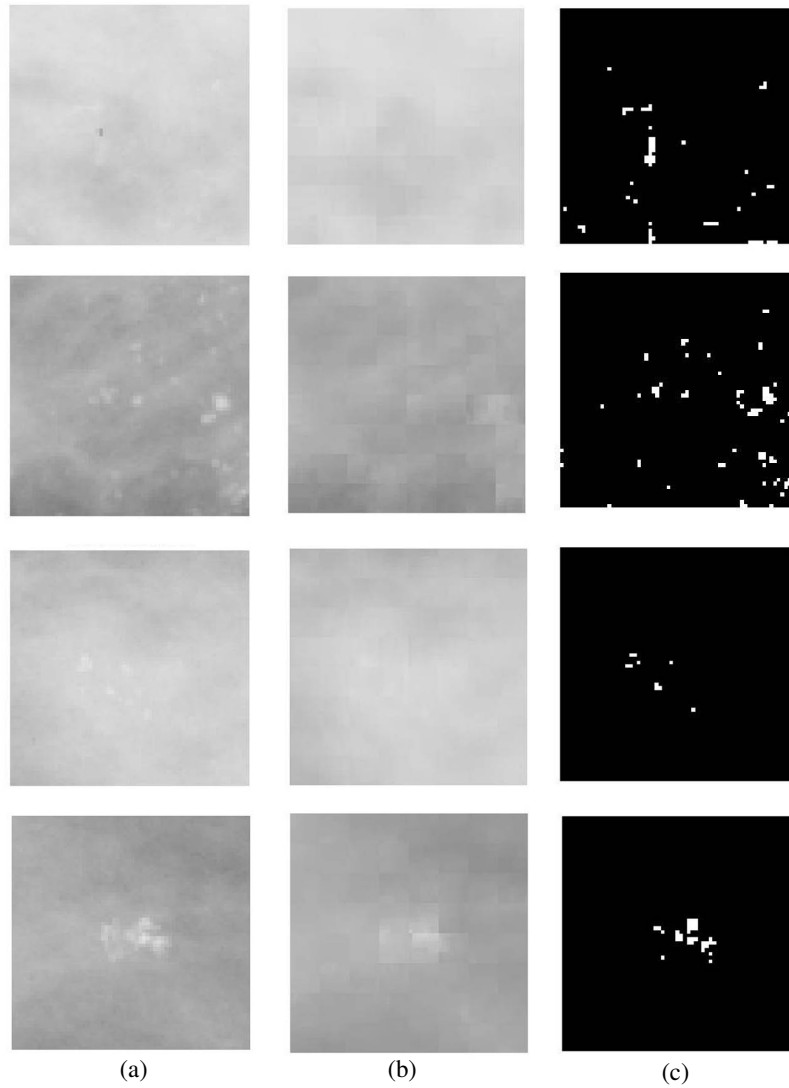


Fig.5.23 (a) Original Mammogram (b) – (c) Modeled Mammogram and Detected Microcalcifications using Entropy Method (ROI-64×64, range-8×8)

Table 5.11 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Entropy based Method

Mammo grams	No. of Samples	Entropy Method						
		TP	FP	TN	FN	Sensitivity	Specificity %	Time (minutes)
Normal	346			313	33		90.462	7.786
With Micro calcifications	201	172	29			85.572		7.117
Total	528	172	29	313	33	85.572	90.462	7.452

Table 5.12

Variation in time taken and Average number of domains searched for different range sizes in Entropy based method (ROI 64×64)

Size of range	Time Taken (minutes)	No. of Domain searched /range
2×2	71.806	768.403
4×4	13.02	412.355
8×8	7.452	385.674
16×16	2.546	101.813
32×32	1.5765	101.81

5.10.1.5. Implementation of Mass Center Method

The original image used for modeling, starting image used for decoding and the modeled image produced in every consecutive iteration are shown in fig.24 (a) - (o) respectively. Up to the ninth iterations, clear blocking effects were present in the image. It starts converging after 12 iterations. The black binary image obtained after the subtraction of the modeled image from the original mammogram is illustrated in fig. 24 (p).

The fig.5.25 shows the original and modeled images for normal mammograms using the mass center method. The plot of MSE correlation and PSNR obtained after each iteration are shown in the fig. 5.26(a)-(c).

The parameters obtained between the original and modeled mammograms after convergence have been achieved is shown in table 5.13.

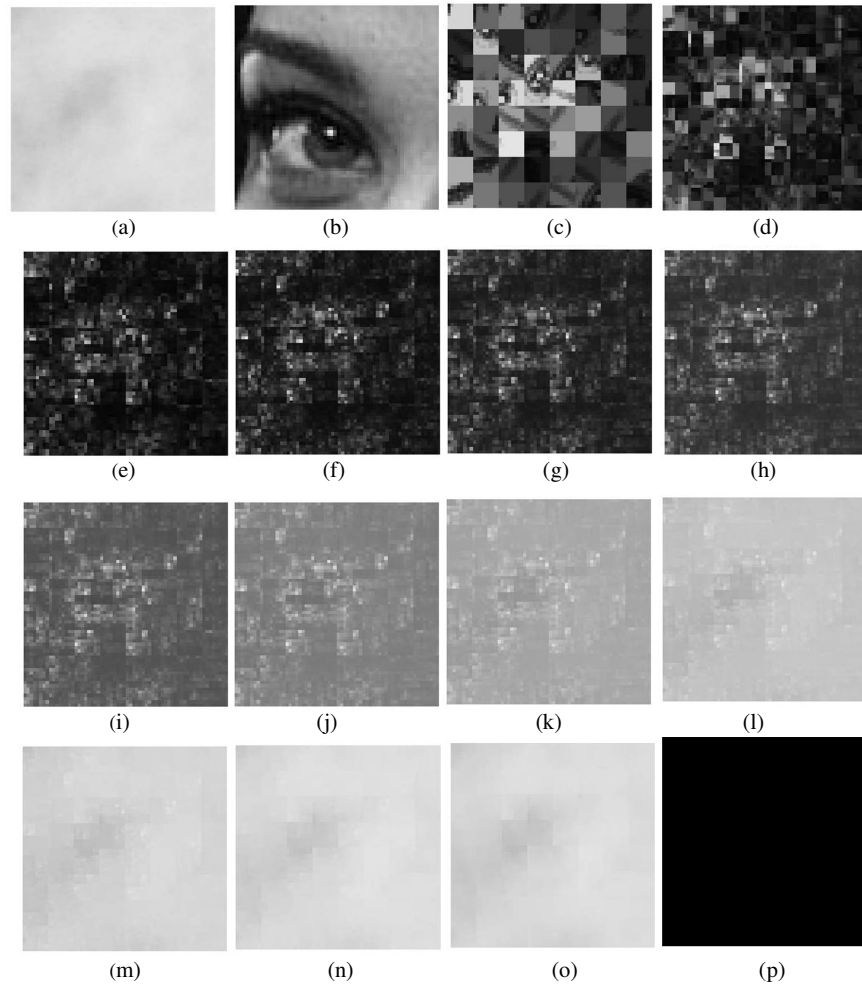


Fig 5.24 Modeling by Mass Center Method (ROI-64×64, range-8×8) (a) Original Image to be modeled (b) Arbitrary chosen Starting Image (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)

There were no deviation in the MSE, correlation and PSNR values once the convergence was attained.

After convergence, the modeled image was subtracted from the original image for enhancing the presence of microcalcifications. The difference image was thresholded as mentioned in section 5.5.2. The detection accuracy obtained is 85.572% and is shown in table 5.14 with an encoding time of 1.903 minutes.

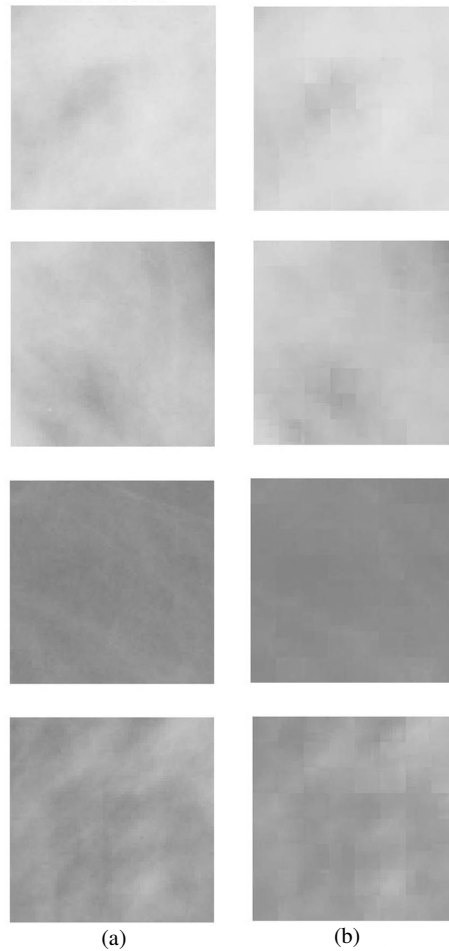


Fig.5.25 (a) Original Mammogram (b) Modeled Mammogram by mass center method(ROI-64×64, range-8×8)

Table 5.15 shows that for a range size of 2×2, the time taken for encoding the range blocks is about one and half hour (90.4422 minutes). On an average of 1482

domain blocks were to be searched per range block. When the range size was increased to 4×4 , the time taken was reduced by a factor of 6.088, i.e. 14.854 minutes. For a range of 8×8 , time required for encoding became 1.5848 minutes. This range size gave a microcalcification detection accuracy of 85.5721%. For range sizes of 16×16 and above, visible blocking artifacts were present in the modeled image, hindering the identification of microcalcifications.

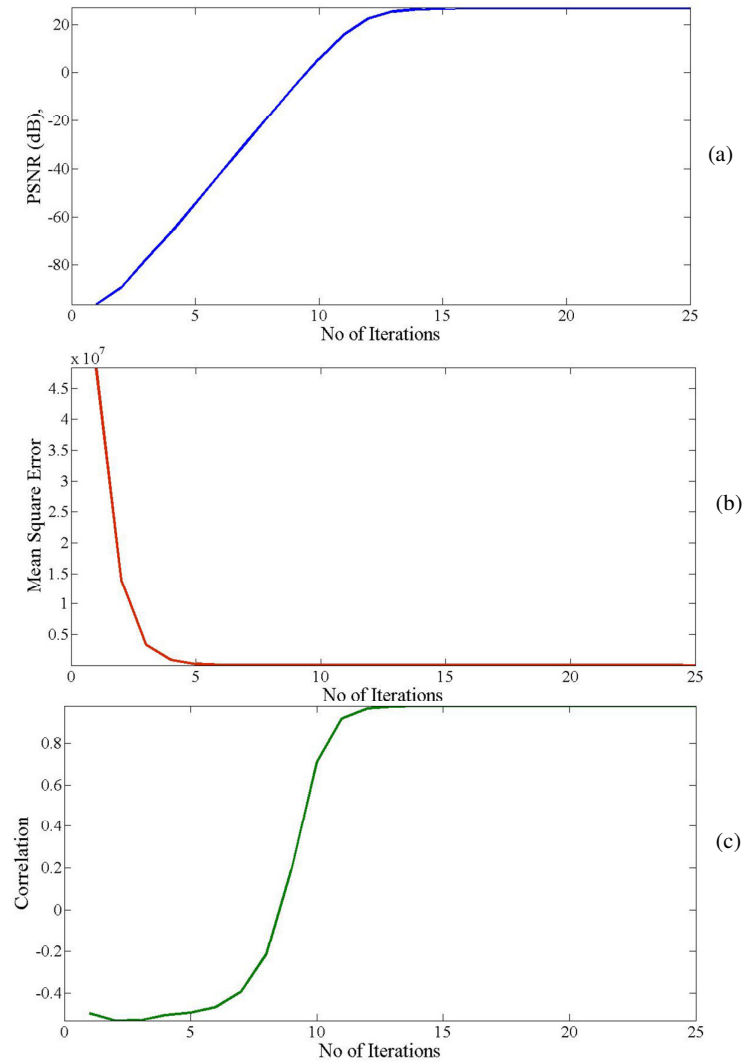


Fig. 5.26 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Mass Center coding method with ROI 64×64 , range size 8×8 .

Table 5.13 Evaluation of the Modeled image obtained using Mass Center fractal modeling method with range size 8×8

Mammogram	MSE	Correlation	PSNR(dB)
Normal	3.1375	0.9710	24.8431

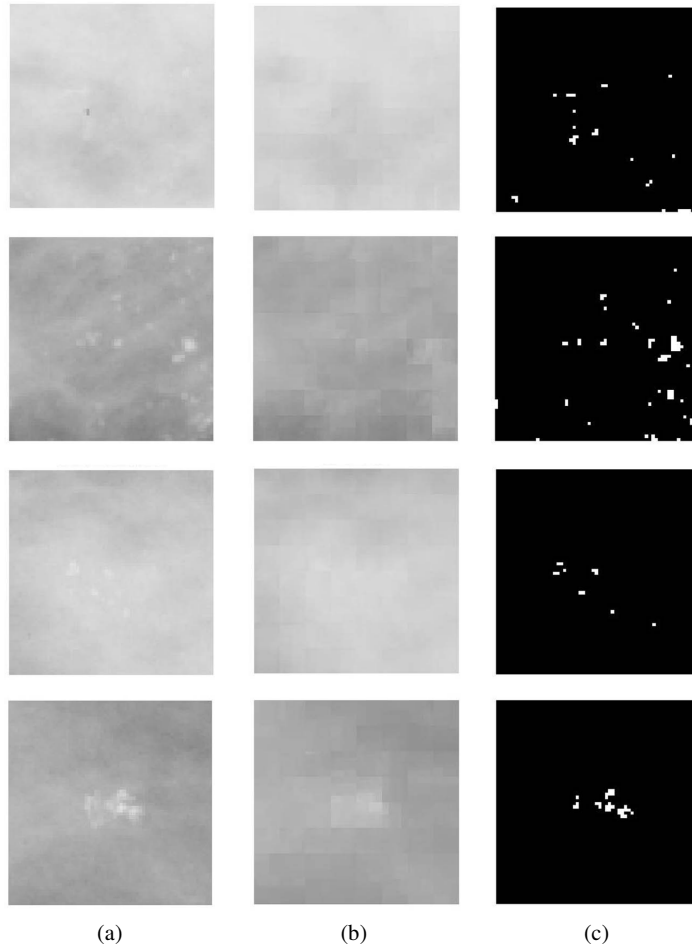


Fig 5.27 (a) Original Mammogram (b) & (c) Modeled Mammogram and Detected Microcalcifications using mass center feature method respectively(ROI- 64×64 , range- 8×8) .

Table 5.14 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Mass Center feature Method with range size 8×8, and ROI 64×64

Mammo grams	No. of Samples	Mass Center Method						
		TP	FP	TN	FN	Sensitivity %	Specificity %	Time (minutes)
Normal	346	-	-	316	30	-	91.329	1.903
With Micro calcifications	201	172	29	-	-	85.5721		1.259
Total	547	172	29	316	30	85.5721	91.329	1.584

Table 5.15 Time taken and Average number of domains searched for different range sizes in Mass center method (ROI 64×64)

Size of range	Time Taken (minutes)	No. of Domain searched /range
2×2	90.4422	1482
4×4	14.854	580.004
8×8	1.5848	352.9
16×16	0.1630	41.375

To further reduce the time required for encoding, shade non shade method of fractal modeling was attempted.

5.10.1.6. Implementation of Shade Non shade Method

The original image, starting image used in decoding, the modeled image obtained after different iterations and the difference image obtained after subtracting the modeled image from the original image are given in fig. 2.28. Fig.5.29 shows the normal mammograms modeled by shade non shade method. The plot of the variation in PSNR, MSE and correlation values with iterations are shown in fig. 5.30. The average value of correlation, PSNR and MSE obtained after each iteration are shown in table 5.16

For the ROIs of 64×64, fractal modeling based on the dynamic range, there were 1022 shade blocks and only two blocks were non shade blocks. As discussed in

section 5.6.4 only the non shade blocks need to be coded by fractal method. For coding the shade blocks, only the mean of those range blocks are required. The modeled image was found to converge to the original image after three to four iterations. The MSE, correlation and PSNR values does not change, even if the iterations are repeated infinite number of times.

The microcalcifications which are detected after subtracting the modeled image from the original mammogram with microcalcification are presented in fig.5.31.

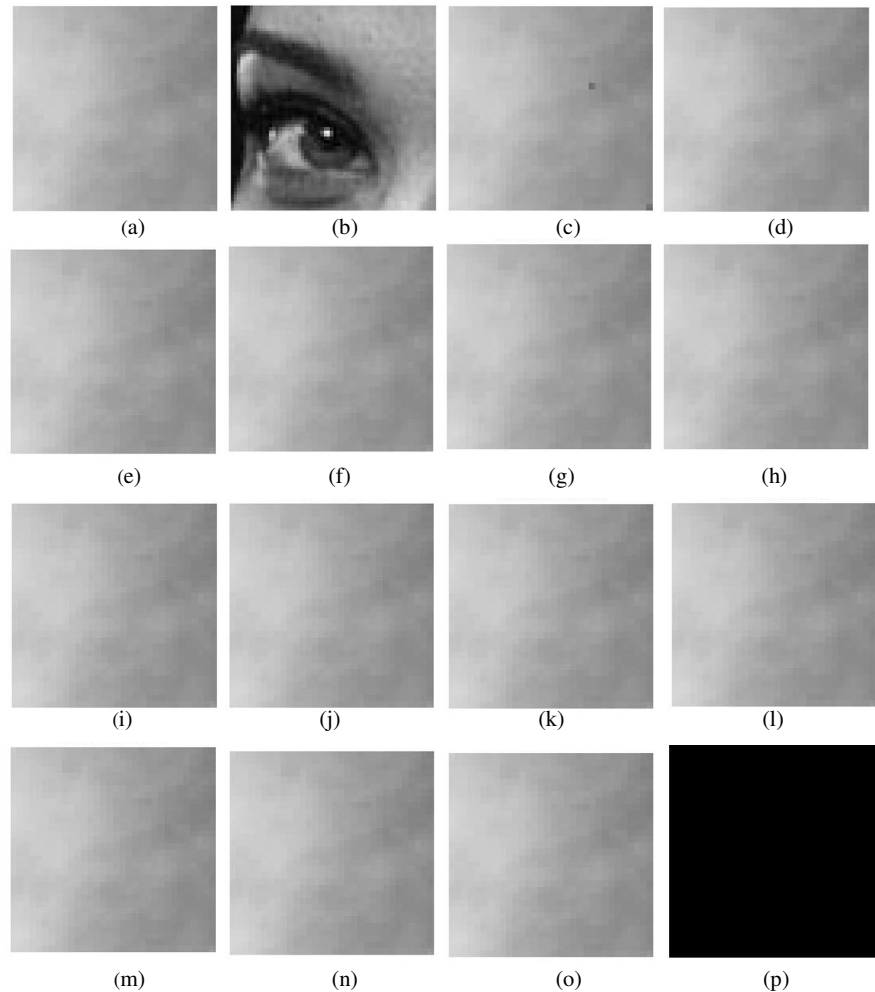


Fig. 5.28 Modeling by Shade Non Shade Method (ROI-64×64, range-8×8) (a) Original Image to be modeled (b) Arbitrary chosen Starting Image (c)-(o) Modeled Image obtained after each iteration (p) Difference image obtained by subtracting (o) from (a)

When the range size was increased, there were more non shade blocks. This is because when the size of range increases, more mammogram features are included in the ROI. Because of this, the dynamic range of its pixels will vary, or increase, causing more blocks to be included in to the non shade category. Hence, when the range size is increased, there are more blocks which are to be modeled by fractal coding method. As a result, the time taken for mammogram coding increases and is shown in table 5.18.

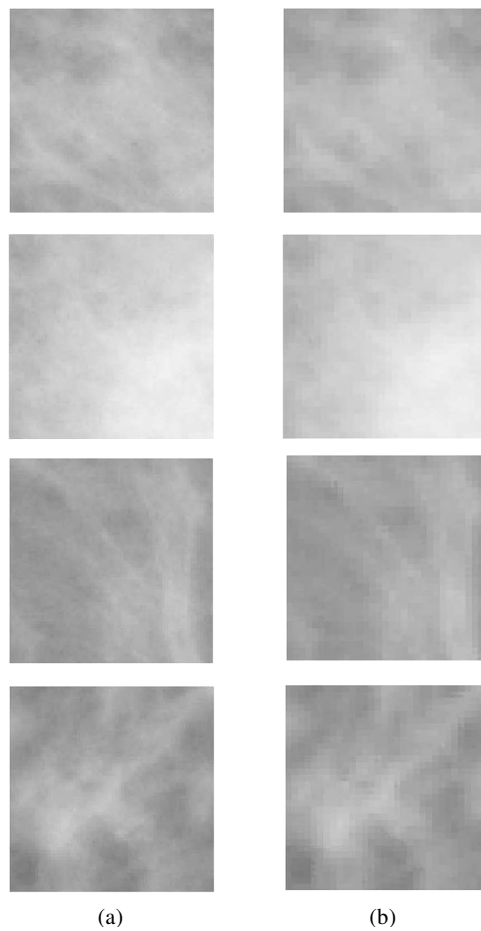


Fig. 5.29(a) Original Mammogram (b) Modeled Mammogram using Shade Non shade method (ROI 64×64, range size 2×2)

Also as the range size increased there were blocking artifacts present in the modeled image. This produced false detection. Thus the optimum size of range was chosen as 2×2 , which gave the highest detection accuracy of 91.54% with the least time of 0.275 minutes taken for encoding mammograms with microcalcifications.

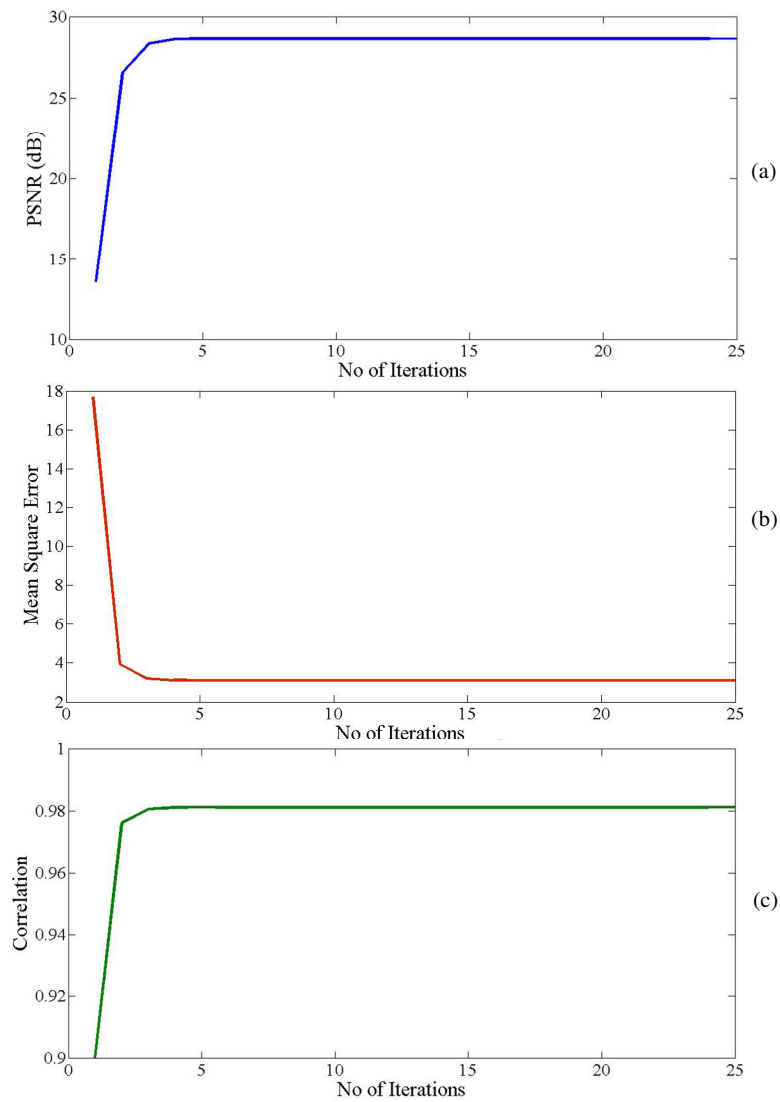


Fig. 5.30 (a) PSNR, (b) MSE and (c) Correlation of original and modeled image using Shade Non shade method with ROI 64×64 , range size 2×2 .

Table 5.16 Evaluation of the Modeled image obtained using Shade Non shade fractal modeling method for a range size of 2×2

Mammogram	MSE	Correlation	PSNR(dB)
Normal	2.2102	0.9974	45.7108

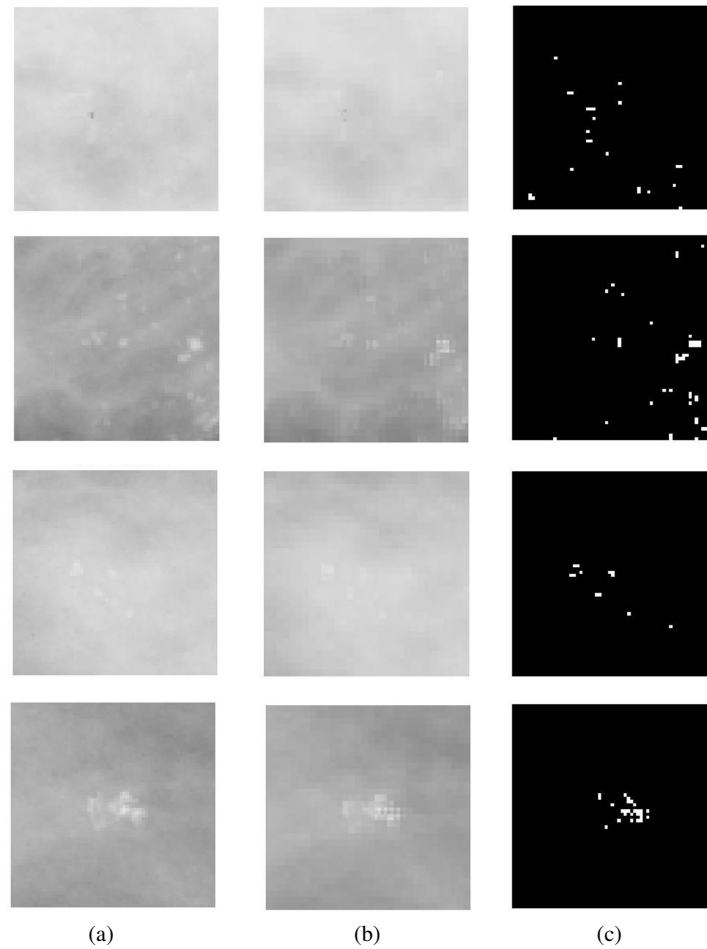


Fig.5.31 (a) Original Mammogram (b) & (c) Modeled Mammogram and Detected Microcalcifications using Shade non shade method respectively(ROI-64×64, range-2×2).

Table 5.17 Detection Sensitivity, Specificity, Average Number of Domains searched and Encoding time for Shade Non shade block Method

Mammo grams	No. of Samples	Shade Nonshade Method						
		TP	FN	TN	TP	Sensitivity %	Specificity %	Time (minutes)
Normal	346			328	18		94.796	0.212
With Micro calcifications	201	184	17			91.542		0.275
Total	528	177	24	328	18	91.542	94.796	0.244

Table 5.18 Time taken, average number of domains searched and the Detection Accuracy for different range sizes in shade non shade method (ROI 64×64)

Size of range	Time Taken (minutes)	No. of Shade Blocks	No. of Non Shade Blocks	Avg. No. of Domain searched	Detection Accuracy (%)
2×2	0.244	1022	2	12	91.542
4×4	1.0112	231	25	760.76	89.635
8×8	3.3547	22	490	1591.2	85.231

5.11 Comparison of the different fractal modeling methods

Results of the above methods are compared next. As conventional fractal modeling method resulted in a poor diagnostic accuracy, it is not used for the comparison.

The time taken for encoding, for various range sizes in different methods, is compared initially. Fig. 5.32 illustrates the variation in the time taken for fractal image modeling for different range sizes. Different range sizes considered were 2×2, 4×4, 8×8, 16×16 and 32×32. The figure shows the time taken for an ROI of size 64×64. For a range of 2×2, the number of domains searched was very large for all the methods. Since modified conventional method is searching the entire image and is considering the entire domain block; the time for encoding was 315 minutes. But as the range size was increased, the number of domains included in the domain search was reduced in the chosen ROI, and therefore the time taken was also reduced. The

same pattern is obtained with the entropy and mass centre method, but with the time much less for the range size 2×2 .

For shade and non shade block method, for a range size of 2×2 , 1022 blocks are shade blocks, on an average only the remaining 2 blocks need to be coded by fractal method. Therefore the time taken is very less for this method. But as the range size increases, more blocks will be included into the shade block category, and hence it requires more time for encoding as more blocks are to be coded by the fractal method. Here, time increases as the range size is increased and therefore is not included in the plot. But as mentioned in section 5.10.1.6, the optimum range size was 2×2 with an average encoding time of 0.244 minutes.

Fig.5.33 summarizes the time taken to encode the mammogram using modified conventional, mean variance, entropy, mass center and shade and non shade fractal modeling methods, using bar graph representations.

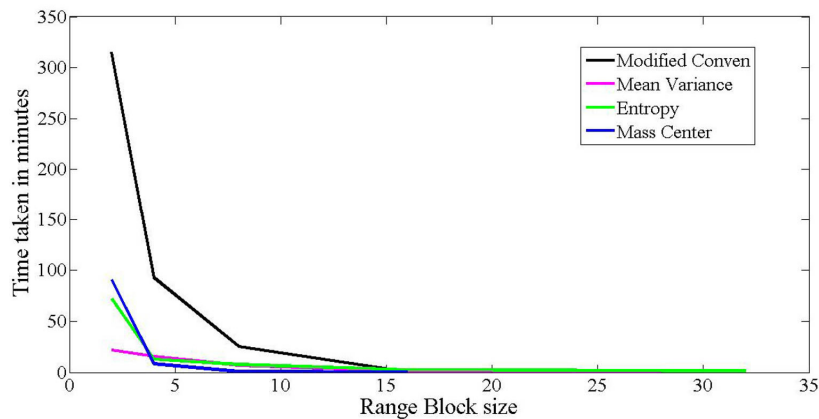


Fig. 5.32 Comparison of the variation in time for the Modified conventional, Mean variance, mass center and entropy methods for different block sizes

It is evident that, compared to the modified conventional fractal coding method, the time taken for other methods are less. The average time for encoding are 26.241minutes, 8.399 minutes, 7.452 minutes, 1.584minutes, and 0.244minutes for modified conventional, mean variance, entropy ,mass center and shade non shade

methods respectively. The minimum time was taken by the shade non shade method. This is because only non shade blocks need to be modeled by fractal method and number of such blocks are very less.

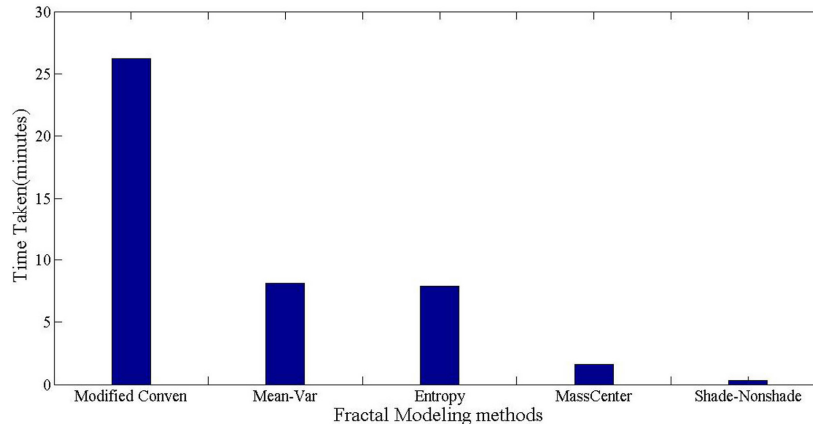


Fig. 5.33 Comparison of the time taken for encoding for modified fractal method and mean variance, entropy, mass center and shade non shade methods

Figs.5.34-5.36 depicts the variation in the PSNR, MSE and correlation between the original and the corresponding modeled image, after every iteration for the above methods.

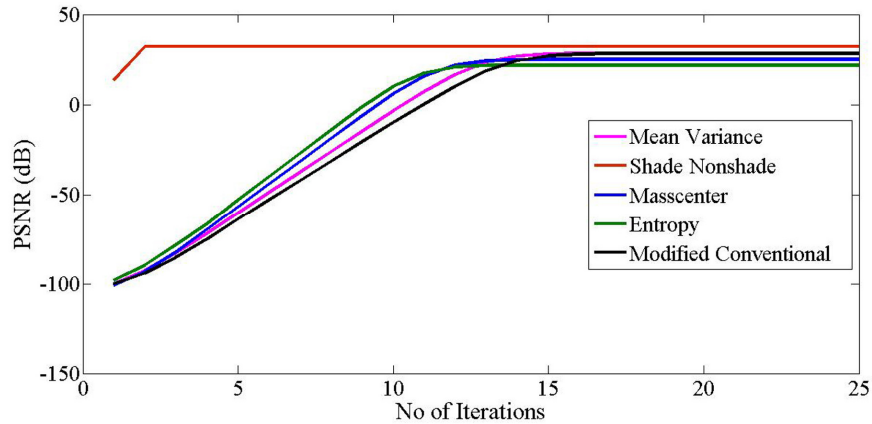


Fig 5.34 Comparison of the PSNR between the modeled image and the original image for the Modified Conventional, Mean Variance, Entropy, Mass Center and Shade-Non Shade Blocks methods of fractal Modeling

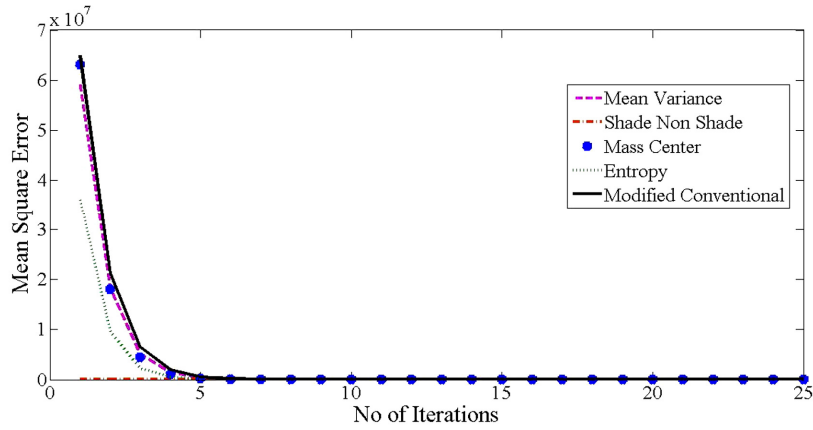


Fig 5.35 Comparison of the Mean Square Error between the modeled image and the original image for the Modified Conventional, Mean Variance, Entropy, Mass Center and Shade-Non Shade Blocks methods of fractal Modeling

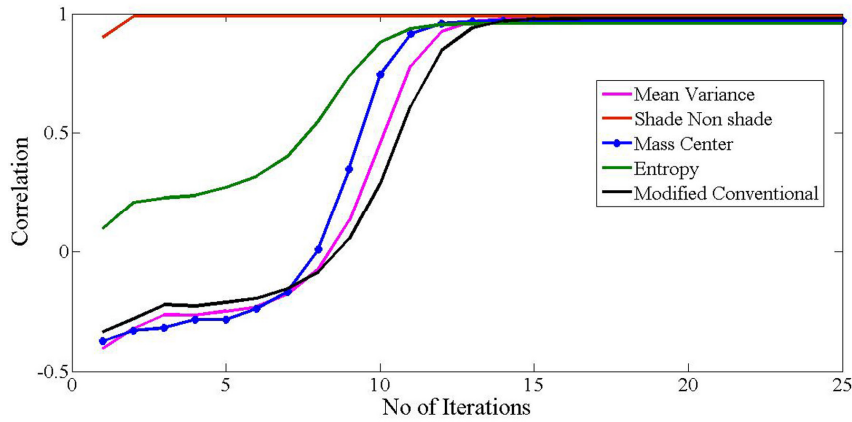


Fig 5.36 Comparison of the Correlation between the modeled image and the original image for the Modified Conventional, Mean Variance, Entropy, Mass Center and Shade-Non Shade Blocks methods of fractal Modeling

The PSNR, MSE and correlation between the original and the modeled image will become a constant value after 12 iterations for all the methods except shade and non shade method. In shade and non shade method, the final modeled image is obtained after 2 iterations as shown in fig. 5.28. Therefore, the final stable values for these parameters are attained in this method in lesser iterations.

Table 5.19 shows the total number of mammograms (both normal and mammograms with microcalcifications) modeled by the different fractal modeling methods and the corresponding detection accuracies obtained. The highest detection accuracy was obtained for the shade and non shade method with the minimum encoding time.

Table 5.19 Comparison of the different microcalcification detection methods (ROI 64× 64, range 8x8, for shade non shade: range 2×2)

Mammo grams	No.of Samples	Methods	TP	FP	TN	FN	% Detection	Time for encoding (minutes)
Normal	346	Conventional			326	20	94.219	26.756
		Mean-Variance			326	20	94.219	8.59
		Entropy			313	33	90.462	7.786
		Mass Center			316	30	91.329	1.903
		Shade Non shade			328	18	94.796	0.2121
With Microcalcifications	201	Conventional	166	35			82.587	25.726
		Mean Variance	166	35			82.587	8.207
		Entropy	172	29			85.572	7.117
		Mass center	172	29			85.572	1.259
		Shade Non shade	184	17			91.542	0.275

The detection accuracy obtained using different fractal modeling methods, for mammograms with microcalcification alone are summarized in table 5.20. The number of domains searched per range block and the average number of ranges coded using fractal encoding method are also illustrated in this table. There are less number of ranges to be encoded by fractal method for shade non shade method when compared to other methods. Therefore, this method gave the highest microcalcification detection accuracy of 91.542% with the least amount of time of 0.275 minutes.

Thus for the normal mammograms and the mammograms with microcalcifications the highest performance is obtained for the shade and non shade method.

Table 5.20

Comparison of Microcalcification Detection %, Average Fractal encoding time, Average number of domains searched and the Average No. of Ranges coded for different microcalcification detection methods(ROI 64× 64 range 8x8.For shade non shade: range 2×2)

Method	Detection %	Average Encoding Time(minutes)	Average No. of Domains Searched	Average No. of Ranges Coded
Modified Conven	82.587	25.726	2013	72
Mean Variance	82.587	8.207	947.46	71
Entropy	85.572	7.117	385.674	72
Mass Center	85.572	1.259	362.9	66
Shade Non Shade	91.542	0.275	13	2

5.12 Chapter Summary

In this chapter, four different fast fractal modeling methods for improving the microcalcification detection accuracy in mammograms are introduced. In the conventional method of fractal image modeling, the best matching domain for each range block is searched after leaving a gap of R or R/2 pixels from the current

domain. To prevent the missing of the smallest microcalcifications of one pixel size, the domain search is to be started from the next adjacent pixel onwards. But, since microcalcifications are occurring as single or isolated small clusters in a mammogram, when the modeled image is subtracted from the original image the presence of microcalcifications should be enhanced. For this reason, the best matching domain for range block should not contain the region corresponding to the range block. This led to the modified fractal image modeling method. Here the microcalcification detection accuracy is 82% but the time taken to encode the mammogram is very large of 26.241 minutes when compared to the conventional modeling method, with detection accuracy and time for encoding as 46.268% and 0.7616 minutes respectively.

To reduce the time taken for modeling, instead of checking for the complete domain pool, domain search based on the mean-variance, entropy, mass center and shade and non shade regions were introduced. Out of these methods, the shade and non shade block method was found to be the fastest, which took only 0.244 minutes for modeling and also produced the highest detection accuracy of 91.52%.

Microcalcifications are early indications of breast cancer. The methods presented in this chapter provide a fast and efficient method to identify the presence of these microcalcifications.

This method can be used for the development of a new software package, which can be used for the first screening of mammogram. The software can aid a radiologist when large number of mammograms is to be screened.

Chapter 6

Conclusions and Future Scope

The main conclusions of this research and the scope for further work are presented in this chapter.

6.1 Thesis Highlights

This chapter brings the thesis to a close by presenting the conclusions drawn from the research. The topic of research was introduced in chapter one. A medical perspective of breast cancer was examined in chapter two. A brief overview of fractals was given in chapter three. The two major contributions of the research are presented in chapters four and five. In chapter four, the different methods for estimating fractal dimension and the different features which were derived from these fractal dimensions for the classification of mammograms are detailed. Chapter five is dedicated to the modeling of mammograms using fractal method, for the identification of microcalcifications in mammograms.

6.2 Classification of mammograms by fractal features

Different fractal features derived from fractal dimension were developed, for the classification of mammograms into normal and abnormal. The abnormalities that are considered in this research are masses and microcalcifications. Also these abnormalities can be malignant and benign.

The classification problem was approached initially by considering only the feature, fractal dimension (FD). Differential Box Counting (DBC); Blanket and Triangular Prism Surface Area (TPSA) methods were the different fractal dimension estimation methods used for finding the fractal dimension. It is seen in chapter 4 (section 4.9.2) a classification accuracy of 80.17% was obtained using the TPSA method. Both the blanket and DBC methods gave an accuracy of 6.436% only. The box plots illustrated the overlap of the FD values between the different classes. The Receiver Operating Characteristics (ROC) analysis was done to evaluate the performance of different fractal dimension estimation methods in the categorization of mammograms. Thus, it was evident that even though fractal dimension can represent irregularity of a surface, it cannot completely categorize mammograms into the different classes. Therefore, six fractal features derived from the above three fractal dimension estimation methods, were considered for better classification.

The fractal feature f_1 to f_5 are already used in literature for segmentation. The sixth feature f_6 was developed during course of this research. The results were analyzed using ROC. The best results were obtained with the feature f_6 computed using the Triangular Prism Surface Area method. The table 4.25 gives the number of mammograms which were correctly classified by the feature f_6 . It is observed that from the total of 799 mammograms, except 11 benign microcalcifications, all the mammograms were correctly classified by this feature f_6 obtained using TPSA method. This gave a total classification accuracy of 98.925%. But the same feature obtained by blanket and DBCM gave an accuracy of only 30.413% and 23.905% respectively.

The discriminative nature of the feature f_6 was validated by carrying out Statistical Analysis. The performance of the fractal features were compared with the conventional textural features in classifying the mammograms. But, none of the textural features could effectively classify mammograms as efficiently as fractal feature f_6 . Classification accuracy and Area Under ROC (AUC) are used for comparing of their efficiency.

6.3 Detection of Microcalcifications by fractal modeling

Microcalcifications are of very small size and are camouflaged in the breast tissues making them difficult to detect even by an experienced radiologist. Therefore, in this research, to identify the presence of microcalcifications, mammograms were modeled by fractal modeling technique. The key challenge in any fractal image modeling scheme is the enormous amount of time taken during the encoding process. In conventional fractal modeling method, during encoding, the matching domain is searched for each range by leaving a gap of R or $R/2$ pixels from the present domain. But, as smaller microcalcifications of the size of one pixel will be missed by this procedure, the conventional method was modified by searching the matching domain from the next adjacent pixel itself, of the current domain. This was named as the modified conventional fractal modeling method. This increased the time taken for encoding by a factor of 34.988, but, the microcalcification detection accuracy was

significantly increased from 46% of the conventional method to 82%. The overall detection accuracy (for both normal and mammograms with microcalcifications) for modified fractal modeling was obtained as 88.4035% when compared to the 67% of the conventional fractal modeling method. Since, the time taken for this modified method was very high, different methods were tried to reduce the encoding time.

Instead of searching all the domain blocks, only those domains which have the same properties of mean variance, entropy, mass center and dynamic range need to be searched for a particular range block. The mean variance, entropy and mass center methods reduced the average encoding time to 8.117minutes, 7.898minutes and 1.5848 minutes respectively. The shade and non shade method which divided the domains based on dynamic range reduced the time to 0.2937 minutes and also gave the highest microcalcification detection accuracy of 91.542%. This is due to the fact that shade blocks are those blocks which do not have significant variations in the dynamic range of the pixels. Most of the blocks in the mammograms will be shade blocks, as microcalcifications are existing as single or as isolated clusters and are not covering the entire mammogram region. Therefore, the average number of domains to be encoded by fractal method dropped by a factor of 5.143, with respect to the other methods.

Thus, the developed fractal modeling method could effectively detect the presence of microcalcifications in lesser time without compromising the detection accuracy. These results are published in (San 2010).

6.4 Suggestions for Future research

Although the present research gave good results, certain proposals for future work are listed below:

- The key point in the survival of the patient is the stage at which cancer is detected. If sufficient data at different stages of the disease are available, the research can be done for predicting cancer.
- In this research only deterministic properties of fractals were used. Properties like fractional Brownian motion (fBm) may be employed for the

classification and detection. Also, combination of these two can be tried to further reduce the encoding time and increase the classification accuracy.

- If the encoding time can be further reduced, this method can be utilized for real time implementation.
- Research may be done to study the shape of the abnormalities present in mammograms. Fractals provide an excellent means for studying the irregular geometric patterns.
- During fractal modeling of mammograms; only square blocks were considered in this research. Adaptive domain sizes may help reducing the time required for encoding.
- Only spatial domain analysis was done in this research. It may be possible to obtain better results when other transforms are integrated with fractal properties.

Bibliography

- (ACS 2008) Retrieved from American Cancer Society Cancer facts and Figures, 2008.: [http://www.cancer.org/downloads/STT/BCFF_Final .pdf](http://www.cancer.org/downloads/STT/BCFF_Final.pdf)
- (Bac2010) Backes, A. R., & Bruno, O. M. (2010). Shape classification using complex network and Multi-scale Fractal Dimension. *Pattern Recognition Letters* , 31, 44-51.
- (Bae 2002) Baeg, S., & Kehtarnavaz, N. (2002). Classification of Breast Mass Abnormalities using Denseness and Architectural Distortion. *Electronic Letters on Computer Vision and Image Analysis* , 1 (1), 11-20.
- (Bar 1988) Barnsley, M. F. (1988). *Fractals Everywhere*. New York: Academic Press.
- (Bas 2004) Base, A. M. (2004). *Pattern Recognition for Medical Imaging*. London: Elsevier Academic Press.
- (Bea 1990) Beaumont, J. M. (1990). Advances in block based fractal coding of still pictures. *IEE Colloquium on The Application of Fractal Techniques in Image Processing*. IEE Digest No. 1990/171.
- (Bel 2009) Bellenir, K. (2009). *Health Reference Series: Breast Cancer Source book*. Detroit: Omnigraphics.
- (Ber 1997) Berizzi, F., Mese, E. D., & Pinelli, G. (1997, October 14-16). A Two-Dimensional Fractal Model of Evaluation. *Radar 97* , 189-193.
- (Bew 2004) Bewick, V., Cheek, L., & Ball, J. (2004). Statistical Review 13: Receiver Operating Characteristic Curves. *Critical Care* , 8 (6), 508-512.
- (Bic 2010) Bick, U., & Diekmann, F. (2010). *Medical Radiology Diagnostic Imaging Digital Mammography*. Berlin Heidelberg: Springer-Verlag.
- (Boc 2004) Bocchi, L., Coppini, G., Nori, J., & Valli, G. (2004). Detection of Single and clustered Microcalcifications in Mammograms using Fractal models and Neural Networks. *Medical Engg Physics* , 303-312.
- (Bol 2007) Bolotov, V. N., & Tkach, Y. V. (2007). UWB Communication Using Fractal Signals. *Proceedings of the 4th European Radar Conference*, (pp. 393-396). Munich, Germany.
- (Bra 1999) Brake, G. M., & Karssemeijer, N. (1999). Single and Multiscale Detection of Masses in Digital Mammograms. *IEEE Transactions on Medical Imaging* , 18 (7), 628-639.
- (BRC 2008) *Pictures of Breast Anatomy*. (2008, June 24). Retrieved May 25, 2011, from Breastcancer.org: http://www.breastcancer.org/pictures/breast_anatomy/
- (Bre 2011) *Breast Cancer*. (2011). Retrieved August 07, 2011, from Roko Cancer Campaign: <http://www.rokocancer.org/aboutcancer>

-
- (Bus 2002) Bushberg, J. T., Seibert, J. A., Leidholdt, J. E., & Boone, J. M. (2002). *Essential Physics of medical imaging*. Philadelphia: Lippincott Williams & Wilkins.
- (Cab 2002) Cabrelli, C. A., Falsetti, M. C., & Molter, U. M. (2002). Fractal Block-Coding: A Functional Approach for Image and Signal Processing. *Computers and Mathematics with Applications*, 1-18.
- (Cal 1990) Caldwell, C. B., Stapleton, S. J., Holdsworth, D. W., Jong, R. A., Weiser, W. A., Cooke, G., et al. (1990). Characterisation of mammographic parenchymal pattern by fractal dimension. *Journal of Physics in Medicine and Biology*, 35 (2), 235-247.
- (Cha 2009) Chanda, B., & Majumdar, D. D. (2009). *Digital Image Processing and Analysis*. New Delhi: PHI Learning Private Limited.
- (Chg 2004) Chang, R. F., Chen, C. J., Ho, M. F., Chen, D. R., & Moon, W. K. (19-21 May 2004). Breast Ultrasound Image Classification Using Fractal Analysis. *Proc. of the Fourth IEEE Symposium on Bioinformatics and Bioengineering*. Taichung, Taiwan: IEEE.
- (Cht 2004) Chatterjee, A., & Rakshit, A. (2004). Influential Rule Search Scheme (IRSS)—A New Fuzzy Pattern Classifier. *IEEE Transactions on Knowledge and Data Engineering*, 16 (8), 881-893.
- (Chd 1995) Chaudhuri, B. B., & Sarkar, N. (1995). Texture Segmentation Using Fractal Dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17 (1), 72-77.
- (Che 2005) Chen, D. R., Chang, R. F., Chen, C. J., Ho, M. F., Kuo, S. J., Chen, S. T., et al. (2005). Classification of breast ultrasound images using fractal feature. *Journal of Clinical Imaging*, 29, 235-245.
- (Cen 1998) Chen, E.-L., Chung, P. C., Chen, C. L., Tsai, H. M., & Chang, C. I. (1998). An Automatic Diagnostic System for CT Liver Image Classification. *IEEE Transactions on Biomedical Engineering*, 45 (6), 783-794.
- (CeT 1998) Chen, T. (1998, March). The Past, Present and Future of Image and Multidimensional Signal Processing. *IEEE Signal Processing Magazine*, 21-58.
- (Chu 2003) Chung, K. H., Fung, Y. H., & Chan, Y. H. (2003). Image Enlargement using Fractal. *Proceedings of ICASSP*, 6, pp. 273-276.
- (Cla 1986) Clarke, K. C. (1986). Computation of the fractal dimension of topographic surfaces using the Triangular Prism Surface Area method. *Computers and Geosciences*, 12 (5), 713-722.
- (Coc 1996) Cochran, W. O., Hart, J. C., & Flynn, P. J. (1996). Fractal Volume Compression. *IEEE Transactions on Visualization and Computer Graphics*, 2 (4), 313-322.
- (Cra 1997) Crane, R. (1997). *A simplified approach to Image Processing Classical and Modern techniques in C*. New Jersey: Prentice Hall.
- (Dee 1992) Deering, W., & West, B. J. (1992, June). Fractal Physiology. *IEEE Engineering in Medicine and Biology Magazine*, 40-46.

- (Dvs 1998) Davis, G. M. (1998). A Wavelet-Based Analysis of Fractal Image Compression. *IEEE Transactions on Image Processing*, 7 (2), 141-154.
- (Dvn 1996) Davoine, F., Artonini, M., Chassery, J. M., & Barlaud, M. (1996). Fractal Image Compression Based on Delaunay Triangulation and Vector Quantization. *IEEE Transactions on Image Processing*, 5 (2), 338- 346.
- (Den 1989) Dennis, T. J., & Dessipris, N. G. (1989). Fractal modeling in image texture analysis. *IEE Proceedings*, 136 (5), 227-235.
- (Don 2009) Dong, A., & Wang, B. (23-26 August 2009). Feature Selection and Analysis on Mammogram Classification. *Proc. of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, (pp. 731-735). Victoria, British Columbia.
- (Dou 2009) Dougherty, G. (2009). *Digital Image Processing for Medical Applications*. Cambridge University Press.
- (Du 2002) Du, G., & Yeo, T. S. (2002). A Novel Multifractal Estimation Method and Its Application to Remote Image Segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (4), 980-982.
- (Edg 2004) Edgar, G. A. (2004). *Classics on Fractals*. USA: Westview Press.
- (Elm 1994) Elmore, J. G., Wells, C. K., Lee, C. H., Howard, D. H., & Feinstein, A. R. (1994). Variability in Radiologists' Interpretations of Mammograms. *New England Journal of Medicine*, 331, 1493-1499.
- (Elt 2009) Eltoukhy, M. M., FAYE, I., & Samir, B. B. (2009). Using Curvelet Transform to Detect Breast Cancer in Digital Mammogram. *Proc. of the 5th International Colloquium on Signal Processing & Its Applications*, (pp. 340-345).
- (Eug 2006) Eugene, K.-W. C., & Ong, G. H. (2006). A two-pass Improved Encoding Scheme for Fractal Image Compression. *Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation*, (pp. 214-219). Sydney.
- (Fal 2003) Falconer, K. (2003). *Fractal Geometry: Mathematical Foundations and Applications* (2nd ed.). England: John Wiley & Sons Ltd.
- (Faw 2006) Fawcett, T. (2006). An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27, 861-674.
- (Fay 2009) Faye, I., Samir, B. B., & Eltoukhy, M. M. (2009). Digital Mammograms Classification Using a Wavelet Based Feature Extraction Method. *Proc. of the Second International Conference on Computer and Electrical Engineering*, (pp. 318-322). Dubai, UAE.
- (Fek 2000) Fekkai, S., Al-Akaidi, M., & Blackledge, J. (2000). Fractal Dimension Segmentation: Isolated Speech Recognition. *Proceedings of the IEE Seminar on Speech Coding for Algorithms for Radio Channels*, (pp. 4/1-4/5). London, UK.
- (Fer 2001) Ferrari, R. J., Rangayyan, R. M., Desautels, J. E., & Frere, A. F. (2001). Analysis of Asymmetry in Mammograms via Directional Filtering With Gabor Wavelets. *IEEE Transactions on Medical Imaging*, 20 (9), 953- 964.

- (Fis 1994) Fisher, Y. (1994). *Fractal Image Compression- Theory and Application*. New York: Springer Verlag.
- (Fra 2007) Franceschetti, G., & Riccio, D. (2007). *Scattering, Natural Surfaces and Fractals*. Academic Press Elsevier.
- (Gal 1969) Gallager, S. H., & Martin, J. E. (1969, December). Early Phases in the Development of Breast Cancer. *Cancer* , 1170-1178.
- (Gi 2003) Gi, Y., Wu, Zhi, M., Huang, Ling, Y., & Wen. (2003). Fractal Image compression with Variance and mean. *Proc.of IEEE ICMA, 1*, pp. 353-356.
- (Gia 1999) Giacomazzi, E., Bruno, C., & Favini, B. (1999). Fractal modelling of turbulent mixing. *Combustion Theory Modelling* , 3, 637-655.
- (Gig 2001) Giger, M. L., Karssemeijer, N., & Armato, S. G. (2001). Computer- Aided Diagnosis in Medical Imaging. *IEEE Transactions on Medical Imaging* , 20 (12), 1205-1208.
- (Gon 2005) Gonzales, R. C., & Woods, R. E. (2005). *Digital Image Processing*. Pearson Education .
- (Had 2009) Hadjileontiadis, L. J. (2009). A Texture-Based Classification of Crackles and Squawks Using Lacunarity. *IEEE Transactions on Biomedical Engineering* , 56 (3), 718-732.
- (Ham 1996) Hamzaoui, R. (1996). Decoding algorithm for fractal image compression. *Electronics Letters* , 32 (14), 1273-1274.
- (Han 1982) Hanely, J. A., & McNeil, B. J. (1982). The meaning and Use of the Area Under a receiver Operating Characteristics (ROC) Curve. *Radiology* , 143, 29-36.
- (Hqg 1998) Hanqiang, C., Guangxi, Z., Yaoting, Z., & Zhengbing, Z. (1998). Edge Detection Based On Fractal Coding. *Proc. of the Int'l Conference on Signal Processing*, 2, pp. 1009-1012.
- (Har 1979) Haralick, R. M. (1979). Statistical and Structural Approaches to Texture. *Proceedings of the IEEE* , 67 (4), 786-804.
- (Har 1973) Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for image Classification. *IEEE Transactions on Systems, man and Cybernetics* , SMC-3, 610-621.
- (Hrt 1996) Hart, J. C. (1996). Fractal Image Compression and Recurrent Iterated Function Systems. *IEEE Computer Graphics and Applications* , 25-33.
- (Has 2005) Hassaballah, M., Makky, M. M., & Mahdy, Y. B. (2005). A Fast Fractal Image Compression Method based Entropy. *Electronic Letters on Computer Vision and Image Analysis* , 5 (1), 30-40.
- (Hat 2001) Hatanaka, Y., Hara, T., Fujita, H., Kasai, S., Endo, T., & Iwase, T. (2001). Development of an Automated Method for Detecting Mammographic Masses With a Partial Loss of Region. *IEEE Transactions on Medical Imaging* , 20 (12), 1209-1214.
- (Hea 2001) Heath, M., Bowyer, K., Kopans, D., Moore, R., & Kegelmeyer Jr., P. (2001). *Proceedings of the Fifth International Workshop on Digital Mammography*. Medical Physics Publishing.

- (Hel 2007) Helmborg, G. (2007). *Getting Acquainted with Fractals*. Berlin: Walter de Gruyter.
- (Hub 2009) Huber, P. J., & Ronchetti, E. M. (2009). *Robust Statistics-2nd Edition*. Hoboken, New Jersey: John Wiley and Sons.
- (Hup 2009) Hupse, R., & Karssemeijer, N. (2009). Use of Normal Tissue Context in Computer-Aided Detection of Masses in Mammograms. *IEEE Transactions on Medical Imaging*, 28 (12), 2033-2041.
- (Hur 1993) Hurtgen, B. (1993). Contractivity of fractal transforms for image coding. *Electronics Letters*, 29, 1749-1750.
- (Hur 1994) Hurtgen, B., & Simon, S. F. (1994). On the Problem of Convergence in Fractal Coding Schemes. *Proc. of the IEEE International Conference on Image Processing*, 3, pp. 103-106. Austin, Texas.
- (Ida 1995) Ida, T., & Sambonsugi, Y. (1995). Image Segmentation Using Fractal Coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 5 (6), 567-570.
- (Ift 2003) Iftekharuddin, K. M., Jia, W., & Marsh, R. (2003). Fractal analysis of tumor in brain MR images. *Machine Vision and Applications*, 13, 352-362.
- (Img 1997) *How Mammography is Performed: Imaging and Positioning*. (1997). Retrieved March 01, 2011, from Imaginis.com: <http://www.imaginis.com/breast-health/how-mammography-is-performed-imaging-and-positioning>
- (Jac 1993) Jackson, D. J., & Hannah, S. J. (1993). Comparative Analysis of Image Compression Techniques. *Proceedings of the Twenty-Fifth Southeastern Symposium on System Theory*, (pp. 513-517). Tuscaloosa, AL, USA.
- (Jaq 1993) Jacquin, A. E. (1993). Fractal Image Coding: A review. *Proceedings of the IEEE*, 81, 1451-1465.
- (Jaq 1992) Jacquin, A. E. (1992). Image coding based on a fractal theory of Iterated Contractive Image Transformations. *IEEE Transactions on Image Processing*, 1, 18-30.
- (Jai 1989) Jain, A. K. (1989). *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall Inc.
- (Jan 2006) Jan, J. (2006). *Medical Image Processing, Reconstruction and Restoration - Concepts and Methods*. Boca Raton, FL: CRC Press, Taylor and Francis Group.
- (Jia 1995) Jiang, J. (1995). Image compression with fractals. *Proceedings of IEE Colloquium on Fractals in Signal and Image Processing*, (pp. 7/1-7/3). London, UK.
- (Jng 2009) Jiang, W., Ji, C. C., & Zhu, H. (6-8 Nov. 2009). Fractal Study on Plant Classification and Identification. *Proc. of the International workshop on Chaos-Fractals theories and applications*, (pp. 434-438). China.
- (Jnw 2008) Jianwei, G., & Jinguang, S. (2008). An Image Compression Method of Fractal Based on GSOFM Network. *Congress on Image and Signal Processing*, 421-425.

- (Jor 2006) Jorgensen, P. E. (2006). *Analysis and Probability Wavelets, Signals, Fractals*. New York: Springer
- (Khu 2004) Khuwaja, G. A., & Rezaq, A. A. (2004). Bi-modal breast cancer classification system. *Pattern Analysis and Applications*, 7, 235-242.
- (Kin 2006) Kinsner, W., Cheung, V., Cannons, K., Pear, J., & Martin, T. (2006). Signal Classification Through Multifractal Analysis and Complex Domain Neural Networks. *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 36 (2), 196-203.
- (Kin 2009) Kinsner, W., & Vera, E. (2009). Fractal Modelling of Residues in Linear Predictive Coding of Speech. *Proc. of the IEEE Int'l conf. on Cognitive Informatics*, (pp. 181-187).
- (Kob 1994) Kobatake, H., Yoshinaga, Y., & Murakami, M. (1994). Automatic Detection of Malignant Tumors on Mammogram. *Proc. of the International Conference on Image Processing. 1*, pp. 407-410. IEEE.
- (Kop 2007) Kopans, D. B. (2007). *Breast Imaging, 3rd Edition*. Lippincott Williams & Wilkins.
- (Kub 1988) Kube, P., & Pentland, A. (1988). On the Imaging of Fractal Surfaces. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 10 (5), 704-707.
- (Kuo 2009) Kuo, C. L., Kuo, C. L., Chen, J. L., & Chang, W. D. (2009, March 26-29). Fractal Features for Cardiac Arrhythmias Recognition Using Neural Network Based Classifier. *Proc. of the International Conference on Networking, Sensing and Control* (pp. 930-935). Okayama, Japan: IEEE.
- (Lai 1994) Laine, A. F., Schuler, S., Fan, J., & Huda, W. (1994). Mammographic Feature Enhancement by Multiscale Analysis. *IEEE Transactions on Medical Imaging*, 13 (4), 725-748.
- (Lan 1995) Langi, A., & Kinsner, W. (1995). Consonant Characterization using Correlation Fractal dimension for speech recognition. *Proceedings of IEEE Communications, Power, and Computing Conference*, (pp. 208-213). Winnipeg, Canada.
- (Lev 2010) Levi, D., & Ullman, S. (2010). Learning to classify by ongoing feature selection. *Image and Vision Computing*, 28 (4), 715-723.
- (Li 1997) Li, H., Liu, K. J., & Lo, S. C. (1997). Fractal modeling and Segmentation for the Enhancement of Microcalcifications in Digital Mammograms. *IEEE Transactions on Medical Imaging*, 16 (6), 785-798.
- (LiK 2001) Lin, K. H., Lam, K. M., & Siu, W. C. (2001). Locating the eye in human face images using fractal dimensions. *IEE Proc. Visual Image Signal Processing*, 148 (6), 413-421.
- (Lin 2009) Lin, C. H., Kuo, C. L., Chen, J. L., & Chang, W. D. (2009 March 26-29). Fractal Features for Cardiac Arrhythmias Recognition Using Neural Network Based Classifier. *Proc. of the IEEE International Conference on Networking, Sensing and Control* (pp. 930-935). Japan: IEEE.

- (Liu 2010) Liu, B., Cheng, H. D., Huang, J., Tian, J., Tang, X., & Liu, J. (2010). Fully Automatic and Segmentation-robust Classification of breast tumors based on local texture analysis of ultrasound images. *Pattern Recognition*, 43.
- (Los 2005) Losa, G. A., Merlini, D., Nonnenmacher, T. F., & Weibel, E. R. (2005). *Fractals in Biology and Medicine* (Vol. 4). Basel, Switzerland: Birkhäuser Verlag.
- (Lu 1996) Lu, G. (1996). Motion Compensation Based on IFS. *Proceedings of ICSP*, 2, pp. 918-921. Beijing, China.
- (Luo 2002) Luo, J., Zhu, H., Feng, J., & Yuan, J. (2002). Analysis of 3-d Partial Discharge Patterns in XLPE Power Cable Based on Fractal Theory. *Proceedings of the 2002 IEEE International Symposium on Electrical Insulation*, (pp. 116-118). Boston, MA, USA.
- (Mal 1993) Malassenet, F. J. (1993). Texture Coding using a Pyramid Decomposition. *Proceedings of the IEEE ICASSP*, 5, pp. 353-356.
- (Man 1982) Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. San Francisco, CA: Freeman.
- (Man 1997) Mandelbrot, B. B. (1997). *Fractals in Scaling and Finance- Discontinuity, Concentration, Risk*. New York: Springer .
- (Mao 1998) Mao, F., Zhang, Y., Song, D., Qian, W., & Clarke, L. P. (1998). An Improved Method of Region Grouping for Microcalcification Detection in Digital Mammograms. *Proc. of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. pp. 740-743. IEEE EMB Society.
- (Mar 1979) Martin, J. E., Moskowitz, M., & Milbrath, J. R. (1979). Breast Cancer Missed by Mammography. *American Journal of Radiology*, 132, 737-739.
- (MNT 2004) *Cancer*. (2004). Retrieved February 2011, 27, from Medical News Today: <http://www.medicalnewstoday.com/info/cancer-oncology/>
- (Mtz 2008) Martinez, W. L., & Martinez, A. R. (2008). *Computational Statistics Handbook with Matlab, 2nd edition*. New York: Taylor and Francis Group.
- (Mar 2001) Maryellen L. Giger, N. K. (2001). Computer- Aided Diagnosis in Medical Imaging. *IEEE Transactions on Medical Imaging*, 20 (12), 1205 -1208.
- (Med 2009) *Diseases and Conditions*. (2009, September 24). Retrieved February 24, 2011, from MedicineNet.com: <http://www.medicinenet.com/breast/article.htm>
- (Min 2003a) Mini, M. G. (2003). *Classification of Mammograms and DWT based Detection of Microcalcification*. Kochi: PhD Thesis.
- (Min 2003b) Mini, M. G., & Thomas, T. (2003). A neural network method for mammogram analysis based on statistical features. *Proceedings of TENCON*, (pp. 1489-1492).
- (Mir 1995) Mir, A. H., Hanmandlu, M., & Tandon, S. N. (1995). Texture analysis of CT images. *IEEE Engineering in Medicine and Biology*, 14, 781-786.
- (Mis 2007) Mishra, A. K., Feng, H., & Mulgrew, B. (2007). Fractal Feature based Radar Signal Classification. *Proc. of the International Conference on Radar Systems* (pp. 1-4). Edinburgh, UK: IET.

- (Mit 2001) Mitra, S. K., Murthy, C. A., Kundu, M. K., & Bhattacharya, B. B. (2001). Fractal Image Compression Using Iterated Function System With Probabilities. *Proceedings of the International Conference on Information Technology: Coding and Computing*, (pp. 191-195). Las Vegas, USA.
- (Mol 2001) Molnar, S. M., & Terdik, G. (2001). A General Fractal Model of Internet Traffic. *Proceedings of the 26th Annual IEEE International Conference on Local Computer Networks (LCN'01)*, (pp. 492-499).
- (Mon 1993) Monro, D. M. (1993). Class of fractal transforms. *Electronics Letters*, 29 (4), 362-363.
- (Mon 1992a) Monro, D. M., & Dudbridge, F. (1992). Fractal approximation of image blocks. *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing ICASSP'92*, 3, pp. 485-488.
- (Mon 1992b) Monro, D. M., & Dudbridge, F. (1992). Fractal block coding of images. *Electronic Letters*, 28, 1053-1055.
- (Mon 1994) Monro, D. M., & Woolley, S. J. (1994). Fractal Image Compression Without Searching. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2, pp. 557-560.
- (Mor 2009) Morimoto, T., Nagao, T., Okazaki, K., Nakagawa, Y., & Tangoku, A. (2009). Current status of breast cancer screening in the world. *Breast Cancer*, 16, 2-9.
- (Mud 2000) Mudigonda, N. R., Rangayyan, R. M., & Desautels, J. E. (2000). Gradient and Texture Analysis for the Classification of Mammographic Masses. *IEEE Transactions on Medical Imaging*, 19 (10), 1032-1043.
- (Myi 2005) Myint, S. W., & Lam, N. (2005). A study of lacunarity-based texture analysis approaches to improve urban image Classification. *Computers, Environment and Urban Systems*, 29 (5), 501-523.
- (Nas 2001) Nass, S. J., Henderson, I. C., & Lashof, J. C. (2001). *Mammography and Beyond: Developing Technologies for the Early Detection of Breast Cancer*. Washinton DC: National Acedemy Press.
- (Nem 2009) Nemoto, M., Honmura, S., Shimizu, A., Furukawa, D., Kobatake, H., & Nawano, S. (2009). A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows. *International Journal of Computer Aided Radiology and Surgery*, 4, 27-36.
- (NLM 2010) *Cancer:Medline Plus*. (2010, November 03). Retrieved February 27, 2011, from National Library of Medicine: <http://www.nlm.nih.gov/medlineplus/cancer.html>
- (Obu 2005) Obuchowski, N. A. (2005). Fundamentals of Clinical research for Radiologists. *American Roentgen Ray Society*, 184, 364-372.
- (OSU 2011) *Healthcare Services*. (2011). Retrieved February 24, 2011, from The Ohio State University Medical Center: http://medicalcenter.osu.edu/patientcare/healthcare_services/breast_health/anatomy_of_the_breasts/Pages/index.aspx

- (Oxf 2011) *Oxford Dictionaries*. (2011). Retrieved February 13, 2011, from Oxford University press: http://www.oxforddictionaries.com/view/entry/m_en_gb04_00680#m_en_gb0400680
- (Pan 2010) Pan, S. M., & Lin, C. H. (2010). Fractal Features Classification for Liver Biopsy Images Using Neural Network-based Classifier. *Proc. of the International Symposium on Computer, Communication, Control and Automation*, 2, pp. 227-230.
- (Pau 2005) Paulsen, K. D., Meaney, P. M., & Gilman, L. C. (2005). *Alternative Breast imaging-Four Model based Approaches*. Boston: Keith D. Paulsen, Paul M. Meaney, Larry C. Gilman, AlternatSpringer Science + Business Media Inc.
- (Pei 1988) Peitgen, H. O., & Saupe, D. (1988). *The Science of Fractal Images*. New York: Springer Verlag.
- (Pei 1991) Peitgen, H. O., Jurgens, H., & Saupe, D. (1991). *Fractals for the Classroom- Part one: Introduction to Fractals and Chaos*. New York: Springer Verlag.
- (Pei1992) Peitgen ,Heins-Otto Peitgen, Hartmut Jürgens, Dietmar Saupe, "*Chaos and Fractals- New Frontiers of Science* ", Springer Verlag, New York, 1992.
- (Pel 1984) Peleg, S., Naor, J., Hartley, R., & Avnir, D. (1984). Multiple Resolution Texture Analysis and Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , PAMI-6 (4), 518-523.
- (Pen 2009) Peng, R., Chen, H., & Varshney, P. K. (2009). Noise-Enhanced Detection of Microcalcifications in Digital Mammograms. *Journal of Selected Topics in Signal Processing* , 3 (1), 62- 73.
- (PiH 2006) Pi, M. H., Li, C. H., & Li, H. (2006). A Novel Fractal Image Watermarking. *IEEE Transactions on Multimedia* , 8 (3), 488-499.
- (Pur 2003) Purkait, P., & Chakravorti, S. (2003). Impulse Fault Classification in Transformers by Fractal Analysis. *IEEE Transactions on Dielectrics and Electrical Insulation* , 10 (1), 109-116.
- (PiM 1997) Pi, M., & Ding, J. P. (1997). Fractal Approximation of Image Block in Its Neighbor and Fractal Coding. *Proceedings of the IEEE International Conference on Intelligent Processing Systems* , (pp. 1002-1004.). Beijing,China.
- (Pis 1998) Pisano, E. D. (1998). *Breast Imaging*. Amsterdam: IOS.
- (Pol 1997) Polakowski, W. E., Cournoyer, D. A., Rogers, S. R., DeSimio, M. P., Ruck, D. W., Hoffmeister, J. W, Raines R. A. (1997). Computer-Aided Breast Cancer Detection and Diagnosis of Masses Using Difference of Gaussians and Derivative-Based Feature Saliency. *IEEE Transactions on Medical Imaging* , 16 (6), 811-819.
- (Pov 2000) Polvere, M., & Nappi, M. (2000). Speed up in Fractal Image Coding: Comparison of Methods. *IEEE Transactions on Image processing* , 9 (6), 1002-1009.
- (Pot 1998) Potlapalli, H., (1998). Fractal-Based Classification of Natural Textures. *IEEE Transactions on Industrial Electronics* , 45 (1), 142-150.

- (Rad 2011) *Mammography*. (2011, June). Retrieved July 24, 2011, from Radiologyinfo.org: <http://www.radiologyinfo.org/en/info.cfm?pg=mammo>
- (Rag 2008) Raghav, S., & Mishra, A. K. (19-21 November 2008). Fractal Feature Based ECG Arrhythmia Classification. *Proc. of the IEEE Region 10 Conference TENCN 2008* (pp. 1-5). Hyderabad,India: IEEE.
- (Raj 2009) Rajanna, U., Erol, A., & Bebis, G. (2009). A comparative study on feature extraction for fingerprint classification and performance improvements using rank-level fusion. *Pattern Analysis and Applications* , 13 (3), 263-272.
- (Ran 1997) Rangayyan, R. M., Faramawy, N. M., Desautels, J. L., & Alim, O. A. (1997). Measures of Acutance and Shape for Classification of Breast Tumors. *IEEE Transactions on Medical Imaging* , 16 (6), 799-810.
- (Rin 1994) Rinaldo, R., & Zakhor, A. (1994). Inverse and Approximation Problem for Two-Dimensional Fractal Sets. *IEEE Transactions on Image Processing* , 3 (6), 802-820.
- (Sam 2009) Sameti, M., Ward, R. K., Parkes, J. M., & Palcic, B. (2009). Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer. *IEEE Journal of Selected Topics in Signal Processing, Vol. 3, No. 1, February 2* , 3 (1), 46-52.
- (San 2008a) Sankar, D., & Thomas, T. (2008). Fast Fractal Coding Method for the Detection of Microcalcification in Mammograms. *Proceedings of the IEEE International Conference on Signal Processing, Communications and Networking (ICSCN 2008)*, (pp. 471-476). Chennai.
- (San 2007) Sankar, D., & Thomas, T. (2007). Fractal Modeling of Mammograms based on Mean and Variance for the Detection of Microcalcifications. *Proceedings of the IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'07)* , 2, pp. 334-338. Sivakasi,India.
- (San 2008b) Sankar, D., & Thomas, T. (2008). Locating Microcalcifications in Mammograms using Modified Fractal Coding Method for the Detection of Breast Cancer. *Proceedings of the IETE National Seminar on Information , Communication and Intelligent Systems*, (pp. 59-65). Kochi.
- (San 2010) Sankar, D., & Thomas, T. (2010). A New Fast Fractal Modeling Approach for the Detection of Microcalcifications in Mammograms. *Journal of Digital Imaging* , 23 (5), 538-546.
- (Sar 1994) Sarkar, N., & Chaudhuri, B. B. (1994). An Efficient Differential Box-Counting Approach to Compute Fractal Dimension of Image. *IEEE Transactions on Systems, Man and Cybernetics* , 24 (1), 115-120.
- (Sau 1996) Saupe, D., & Ruhl, M. (1996). Evolutionary Fractal Image Compression. *Proceedings of the IEEE International Conference on Image Processing (ICIP'96)*, (pp. 129-132). Lausanne.
- (Sen 1992) Senevirathne, T. R., Bohez, E. J., & Winden, J. V. (1992). Amplitude Scale Method: New and Efficient Approach to Measure Fractal Dimension of Speech waveforms. *Electronic Letters* , 28 (4), 420-422.
- (Sch 1991) Schroeder, M. (1991). *Fractals, Chaos, Power laws-Minutes from an 218infinite paradise*. New York: W.H. Freeman and Company.

- (Sha 2000) Shang, C., Duly, C., McGrath, J., & Barker, J. (2000). Analysis and Classification of Tissue Section Images using Directional Fractal Dimension Features. *Proc. of the IEEE International Conference on Image Processing, 1*, pp. 164-167.
- (Sic 2007) Sickles, E. A. (2007). Wolfe Mammographic Parenchymal Patterns and Breast Cancer Risk. *American Journal of Roentgenology*, 188, 301-303.
- (Smb 1993) Samarabandu, J., Acharya, R., Hausmann, E., & Allen, K. (1993). Analysis of Bone X-Rays Using Morphological Fractals. *IEEE Transactions on Medical Imaging*, 12 (3), 466-470.
- (Son 2010) Song, D., & Tao, D. (2010). Biologically Inspired Feature Manifold for Scene Classification. *IEEE Transactions on Image Processing*, 19 (1), 174-184.
- (Ste 1997) Stern, I.(1997) On Fractal Modeling in Astrophysics: The Effect of Lacunarity on the Convergence of Algorithms for Scaling Exponents. *Astronomical Data Analysis Software and Systems VI, ASP Conference Series, 125*, pp. 222-225.
- (Stu 2003) Stupnicka, W. S. (2003). *Chaos- Bifurcations and Fractals Around us: A Brief Introduction* (Vol. 47). Singapore: World Scientific Series on Nonlinear science, Series A.
- (Suc 1994) Suckling et.al, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. *International Congress Series*, (pp. 375-378).
- (SzC 2003) SzCkely, N., & Pataki, B. (2003). Detecting Lesions in a Mammogram. *Proc. of the 4th EURASIP Conference on Video/Image Processing and Multimedia Communications*, (pp. 113-118). Zagreb, Croatia.
- (Tab 2003) Tabar, L., & Dean, P. B. (2003). Mammography and breast cancer: the new era. *International Journal of Gynecology and Obstetrics*, 82, 319-326.
- (Tab 2001) Tabar, L., & Dean, P. B. (2001). *Teaching Atlas of Mammography*. Stuttgart: Thieme.
- (Taf 2010) Tafti, P. D., Gonzalo, R. D., & Stalder, A. F. (2010). Fractal Modelling And Analysis Of Flow-Field Images. *Proc. of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, (pp. 49-52). Rotterdam.
- (Tak 2003) Takezawa, M., Haseyama, M., & Kitajima, H. (2003). Ultra Low Bit-Rate Image Coding Algorithm Based on Fractal Image Coding. *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis*, (pp. 1013-1017).
- (TaM 2005) Tang, M., & Wang, H. N. (2005). Feature Analysis of Brain MRI Images Based on Fractal Dimension. *Proceedings of the 27th Annual Conference IEEE Engineering in Medicine and Biology*, (pp. 3245-3248). Shanghai,China.
- (Tan 2009) Tang, J., Liu, X., & Sun, Q. (2009). A Direct Image Contrast Enhancement Algorithm in the Wavelet Domain for Screening Mammograms. *IEEE Journal of Selected Topics in Signal Processing*, 74-80.

- (Tao 2000) Tao, Y., Lam, E. C., & Tang, Y. Y. (2000). Extraction of Fractal Feature for Pattern Recognition. *Proceedings of IEEE International Conference on Pattern Recognition*, 2, 527-530.
- (The 2006) Theodoridis, S., & Koutroumbas, K. (2006). *Pattern Recognition*, 3rd Edition. USA: Academic Press.
- (Tho 1995) Thomas, L., & Deravi, F. (1995). Region-Based Fractal Image Compression Using Heuristic Search. *IEEE Transactions on Image Processing*, 4 (6), 832-838.
- (Tri 1997) Triebel, H. (1997). *Fractals and Spectra-Related to Fourier Analysis and Function Spaces*. Switzerland: Birkhauser Verlag.
- (Vel 1998) Velanovich, V. (1998). Fractal Analysis of Mammographic lesions: A Prospective blinded Trial. *Breast Cancer Research and Treatment*, 49, 245-249.
- (Vin 2003) Vinoy, K. J., Abraham, J. K., & Varadan, V. K. (2003). On the Relationship between Fractal Dimension and the Performance of Multi-Resonant Dipole Antennas Using Koch Curves. *IEEE Transactions on Antennas and Propagation*, 51 (9), 2296-2303.
- (Wei 2005) Weinstein, S., Obuchowski, N. A., & Lieber, M. L. (2005). Fundamentals of Clinical Research for Radiologists : Clinical Evaluation of Diagnostic Tests. *American Journal of Radiology*, 184, 14-19.
- (Wol 1976) Wolfe, J. N. (1976). Breast Patterns as an index of risk for developing breast cancer. *American Journal of Roentgenology*, 26, 1130-1139.
- (Wu 1992) Wu, C. M., Chen, Y. C., & Hsieh, K. S. (1992). Texture Features for Classification of Ultrasonic Liver Images. *IEEE Transactions on Medical Imaging*, 11 (2), 141- 152.
- (Xue 1999) Xuejun, W., Lianyu, C., & Hexin, C. (1999). A Quadtree Classified-based Fractal Image Coding Approach. *Proceedings of the fifth Asia-Pacific Conference on Communication and Fourth Optoelectronics and Communications Conference (APCC/OECC '99)*, 2, pp. 912-915. Beijing, China.
- (Yis 2002) Yisong, C., Guoping, W., & Shihai, D. (2002). Feature Difference Classification Method in Fractal Image Coding. *Proc. of the IEEE ICSP*, (pp. 648-651).
- (Zha 2000) Zhao, D. Y., & Xin, C. H. (2000). An Adaptive Method of IFS Image Compression Based on Creditability. *Proceedings of the fifth IEEE Symposium on Computers and Communications*, (pp. 744-749). France.
- (Zao 2005) Zhao, E., & Liu, D. (2005). Fractal Image Compression Methods: A Review. *Proceedings of the Third International Conference on Information Technology and Applications (ICITA'05), 4-7 July 2005*, pp. 756 – 759, vol.1 China, 1, pp. 756-759. China.
- (Zho 2010) Zhou, Y., Geng, H., Li, M., Yang, Z., & Wang, Z. (2010). Multirange Fractal Approach to Analyze the Microstructure of Sn Melt. *Proceedings of the 3rd International Congress on Image and Signal Processing*, (pp. 3147-3151). Yantai. 220

List of Publications

Journal Papers\Lectures

1. **Deepa Sankar** , Tessamma Thomas, A New Fast Fractal Modeling Approach for the Detection of Microcalcifications in Mammograms, Journal of Digital Imaging, Springer New York, Vol. 23, No- 5, October, 2010, pp 538-546
2. **Deepa Sankar**, Tessamma Thomas, Fractal Features based on Differential Box Counting Method for the Categorization of Digital Mammograms , International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM), Vol-2, 2010, pp.11-19. ISSN: 2150-7988
3. **Deepa Sankar**, Tessamma Thomas, Fractal dimension for the description of Digital mammograms using Triangular Prism Surface Area method, Lectures on Modeling and Simulation- A Trivandrum, India, December 1-3, 2009, Vol. 10 Issue.1 2009, AMSE France, pp. 102-111
4. **Deepa Sankar**, Tessamma Thomas, “Breast Cancer Detection using Entropy based Fractal Modeling of Mammograms International Journal of Recent Trends in Engineering”, Academy Publisher, Finland, Vol. 1, No. 3, May 2009, pp. 171-175
5. Deepa Sankar, Tessamma Thomas, “Classification of Mammograms into Normal, Benign and Malignant based on Fractal Features”, Journal of Medical and Biological Engineering, Taiwan (**Communicated**)

International Conferences

1. **Deepa Sankar**, Tessamma Thomas, “Fractal Signatures for the Characterization of Mammograms”, Proc. of the International Symposium on Innovations in Natural Computing (INC 2009), December 12-13, 2009 at Cochin, India 2009, pp-1463-1468 [**Paper Presented**]
2. **Deepa Sankar**, Tessamma Thomas, “Analysis of Mammograms using Fractal Features”, Proc. of the International Conference on Computer

Information Systems and Industrial Management Applications (CISIM) 2009
December 09-11, 2009, Coimbatore, India, pp-936-941[**Paper Presented**]

3. **Deepa Sankar**, Tessamma Thomas, “Fractal Dimension for the Description of Digital Mammograms using Triangular Prism Surface Area Method”, Proc. of the AMSE International Conference on Modeling and Simulation, MS 09, India – 1-3 Dec, 2009,pp-197-200 [**Paper Presented**]
4. **Deepa Sankar**, Tessamma Thomas, “Analysis of Medical Images using improved Fractal Modeling Approach”, International Conference on Advances in Neurosciences, Dec12-14, 2008, Kochi, pp-59-60
5. **Deepa Sankar**, Tessamma Thomas, “Fast Fractal Coding Method for the Detection of Microcalcification in Mammograms”, Proc. of the IEEE International Conference on Signal Processing, Communications and Networking (ICSCN 2008), Chennai-600044,pp-471-476, January 4-6, 2008 [**Paper Presented**]
6. **Deepa Sankar**, Tessamma Thomas, “Fractal Modeling of Mammograms based on Mean and Variance for the Detection of Microcalcifications”, Proc. of the IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'07), Sivakasi, vol-II, pp-334-338, December 13-15 ,2007 [**Paper Presented**]

National Journal\Conferences

1. **Deepa Sankar**, Tessamma Thomas, “Analysis of Medical Images using improved Fractal Modeling approach”, Journal of Indian Academy of Neurosciences , Vol-15, pp-59-60,2008
2. **Deepa Sankar**, Tessamma Thomas, “Implementation of Mass center Features for the Detection of Microcalcifications” Accepted for Presentation at the National Conference on New frontiers and Current trends in Biotechnology, Madras, India, 18th – 19th October 2008
3. **Deepa Sankar**, Tessamma Thomas, “Locating Microcalcifications in Mammograms using Modified Fractal Coding Method for the Detection of Breast Cancer”, Proc. of the National Seminar on Information , Communication and Intelligent Systems, Kochi, pp-59-65,February 8-9, 2008

Index

A

affine transformation 46, 47,
51,135,140,145
Architectural distortion 19
area under the curve 65, 83, 132
AUC 62, 83, 84, 129, 131, 133, 134, 138
attractor 42, 49, 50,140,141

B

Barnsley 49, 50, 51, 53
benign 3, 10, 14, 16, 17, 18, 20
56, 64, 65, 85, 86, 93, 94, 95, 101, 115,
116,
123, 126, 127, 137, 138, *See*
biopsy 11,56
biopsy 56
Blanket Method 68
box counting method 63, 66, 67
breast 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13,
14,
16, 17, 21
Breast cancer 1, 8, 10, 11, 134, 135, 204

C

CAD 7,65
calcification 13, 16
Cauchy sequence 44, 45
Circumscribed masses 17
Clarke 73
Cluster Prominence 81, 134
Cluster Shade 81, 135
Collage Theorem 49, 136, 140
compression 137
Computed Tomography 61
Contraction Mapping Theorem 47, 48
contractive transformations 136

Contrast 80, 134
convergence 137, 147, 164, 166, 177, 182,
188, 189
Cooper's ligaments 14
Correlation 80, 81, 134

D

DDSM 84, 85, 86, 94, 100, 102, 103, 105,
110, 111, 112, 115, 117, 121, 123, 126,
134,159
Difference entropy 82, 135
Difference variance 82, 135
Differential Box Counting 66,67, 68
Differential Fractal Signatures 72
Dissimilarity 81, 135
Distance Measurement 72
domain blocks 137, 138, 142, 143, 144,
150, 151, 152, 153, 155, 156, 175, 190

E

encoding137, 138, 145, 146, 147, 150, 151,
155, 161, 168, 169, 175, 176, 182, 183,
184, 189, 190, 193, 195, 198, 199, 202,
203, 204
Energy 81, 135
entropy 81,135,151, 153, 182, 183, 199,
204
Euclidean geometry 34, 36, 37

F

fractal34, 35, 36, 37, 38, 39, 43, 44, 49, 50,
51, 52, 53, 54, 55
fractal dimension 36,56, 57, 59, 61, 63,
64, 66, 67, 68, 70, 73, 74, 75, 76, 90, 94,
95, 129, 132, 133, 137

-
- Fractal Feature 56,75, 76, 77, 95, 100, 105,
110, 115, 121
- Fractal Signature 70, 71
- Fractional Brownian Motion 62
- False negative 157
- False Positive 128,157
- fractal image compression 137
- Fractal Modeling 135, 163, 167
- G**
- Georg Cantor 34
- Giuseppe Peano 34
- H**
- Hausdorff 34, 50, 53
- Helge von Koch 34
- Homogeneity 81, 135
- I**
- Image 2, 4, 5
- image acquisition 2, 3, 5, 7
- image enhancement 5, 9
- Information measure of correlation 82,
135
- Inverse difference moment normalized
82, 135
- isometry transformation 144, 147, 152,
154
- Iterated Function Systems (IFS) 48,135,
140
- J**
- Jacquin 136, 155
- Julia 34, 39, 40, 52
- K**
- Kurtosis 78, 79, 134
- L**
- lacunarity 63
- lower surface 69
- M**
- Magnetic Resonance Imaging 7
- malignant 3, 8, 10, 14, 17, 18, 20,56
- mammogram 3, 4, 7, 8, 9, 12, 13,56,134
- Mammography 7,134
- Mandelbrot 34, 40, 52, 53, 55,56,59,70
- mass 56
- mass center 138, 151, 154, 188, 190, 192,
199, 204
- Maximum probability 81, 135
- Mean 78, 88, 90, 97, 102, 107, 111, 117,
124, 134
- Mean and Variance Method 151
- Mean Squared Error 158, 167
- medical imaging 4, 6, 7
- metric space 44, 45, 47, 48, 49, 50
- MIAS 65, 84, 85, 86, 94, 100, 102, 103,
105, 109, 111, 112, 115, 117, 120, 121,
123, 126, 134,159
- microcalcification 138
- Multiple Reduction Copy Machine 40, 41,
42, 43
- P**
- Peak Signal to Noise Ratio 137, 158, 167
- pattern 62, 75
- Q**
- quad treeing 146
- R**
- radiologist 56, 134, 135, 204

-
- range block 138, 142, 146, 150, 152, 153,
156, 162, 175, 181, 183, 184, 190, 203,
204
- Receiver Operating Characteristics 82, 138
- ROC 62, 65, 82, 83, 84, 128, 129, 130, 131,
133, 134, 138
- ROI 159, 161, 163, 164, 166, 167,
169, 173, 174, 175, 176, 179, 181, 183,
185, 186, 187, 191, 192, 193, 194, 195,
196, 198
- S**
- Segmentation 5
- Self similarity 35, 59
- sensitivity 64, 65, 83, 157, 158, 169, 175,
181, 187
- Shade and Nonshade Method 155, 151, 155,
198, 199, 201, 203, 204
- signal to noise ratio 158, 170
- Skewness 78, 134
- specificity 64, 65, 82, 83, 157, 158
- Spiculated lesions 20
- standard deviation 78, 86, 110
- Statistical Analysis 82, 133
- Statistical Descriptors 78
- Sum average 81, 135
- Sum entropy 82, 135
- Sum variance 81, 135
- Symptoms 10
- T**
- texture 61, 63, 65, 75, 76, 78, 79, 80, 134
- Texture features 79
- Triangular Prism Surface Area Method 73
- True Negative 129, 157
- True Positive 128, 157
- tumor 56
- U**
- upper surface 69
- V**
- Variance 64, 78, 82, 135
- W**
- Waclaw Sierpinski 34, 37, 52

Fractal based Techniques for Classification of Mammograms and Identification of Microcalcifications

PhD Thesis

By

Deepa Sankar

Breast cancer has been a leading cause of death among women. The mortality rate of breast cancer can be reduced substantially if detected in the initial stages of the disease. Usually, radiologists rely on computer-aided diagnostic systems for mammographic image interpretations. As mammographic parenchymal and ductal patterns possess fractal properties, fractal analysis can be applied for mammogram analysis effectively.

In this thesis, a new fractal feature was developed for the efficient classification of mammograms into different classes: normal, abnormal-masses (benign and malignant) and microcalcifications (benign and malignant). The presence of microcalcification in mammograms was identified for the early detection of breast cancer. This was done by modeling mammograms using fractals. Also, the newly developed fractal modeling technique is simple, faster and comprehensive.

Department of Electronics
Cochin University of Science and Technology
Kochi-682022. Kerala. India.