MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING

A Novel Approach for Diagnosis of Breast Cancer using Ultrasound Imaging

By

Niamul Quader

Department of Electrical and Electronic Engineering
Islamic University of Technology (IUT)
Gazipur-1704, Bangladesh.

September, 2012.

A Novel Approach for Diagnosis of Breast Cancer using Ultrasound Imaging

A Thesis Presented to the Academic Faculty

by

Niamul Quader
(102606)

In Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE
in
Electrical and Electronic Engineering

Islamic University of Technology
SEPTEMBER 2012

The thesis titled

A Novel Approach for Diagnosis of Breast Cancer using Ultrasound Imaging

Submitted by Niamul Quader, St. No. 102606 of Academic Year 2010-2011 has been found as satisfactory and accepted as partial fulfillment of the requirement for the Degree Master of Science in Electrical and Electronic Engineering.

Board of Examiners

_____
Dr. Kazi Khairul Islam
Professor, EEE Department
IUT, Gazipur - 1704

Chairman
(Supervisor)

_____
Dr. Kaisar Alam
Scientist, Riverside Research, New York, U.S.A.
Visiting Faculty
EEE Department
IUT, Gazipur - 1704

Member
(Co-Supervisor)

_____
Dr. Md. Shahid Ullah
Professor and Head, EEE Department
IUT, Gazipur - 1704

Member
(Ex-Officio)

_____
Dr. Ruhul Amin
Professor, EEE Department
IUT, Gazipur - 1704

Member

_____
Dr. Md. Aynal Haque
Professor, EEE Department
BUET, Dhaka

Member
(External)

Date Approved: 09.09.2012

DECLARATION OF CANDIDATE

It is here by declared that this thesis or any part of it has not been submitted elsewhere for the award of any Degree or Diploma.

Dr. Kazi Khairul Islam, (Supervisor)
Professor, EEE Department
IUT, Gazipur – 1704

Niamul Quader
Student Number: 102606
Academic Year: 2010-2011

Dr. Kaisar Alam, (Co-Supervisor)
Scientist, Riverside Research, New York, U.S.A.
Visiting Faculty Member
EEE Department
IUT, Gazipur - 1704

DEDICATION

This thesis is dedicated to my beloved parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

2D, 3D	Two-dimensional, Three dimensional
GHz	Giga Hertz
BI-RADS	Breast Imaging Reporting and Data System
ROC	Receiver Operating Characteristic
dB	Decibel
C	Convexity
HD	Fractal (Hausdorf) Dimension
FNPA	Four-Neighborhood Pixel Algorithm
stform	Scattered Translates Form
↑	Increase
↓	Decrease
σ	Standard Deviation
ρ	Correlation Coefficient
cov	Co-variance
μ	Mean
NGO	Non-Governmental Organization
WHO	World Health Organization

ACKNOWLEDGEMENTS

First of all I am greatly thankful to the Great Almighty ALLAH (SWT) who has given me the capability to do all the works. I must express my humble gratitude and indebtedness to the thesis supervisor **Prof. Dr. Kazi Khairul Islam,** for his initiative in this field of research, for his valuable guidance, patience and affection throughout this work. His sincere sympathy and kind attitude has been a beacon to my confidence. I must also express my sincere gratitude to the thesis co-supervisor **Dr. Kaisar Alam**, for his affectionate guidance, thorough support, and unfaltering patience and trust in this work. Also, I express our thankfulness to **Prof. Dr. Md. Shahid Ullah**, Head of the Department of Electrical and Electronic Engineering, IUT, Gazipur for providing us with best facilities in the Department, his timely suggestions and critical viewpoints. I would also like to thank **Prof. Dr. Ruhul Amin,** Professor, EEE Department, IUT, **Prof. Dr. Md. Ashraful Hoque**, Professor, EEE Department, IUT, and **Prof. Dr. Aynal Haque,** Professor, EEE Department, BUET, for their kind suggestions in our work.

Last but not least I would like to thank all my friends and well wishers who were involved directly or indirectly in successful completion of this novel work.

NIAMUL QUADER
Student ID: 102606

A Novel Approach for Diagnosis of Breast Cancer using Ultrasound Imaging

Abstract

Breast cancer continues to be the most occurring cancer in women around the globe. With advancements in medical procedures, breast cancers are now virtually treatable in their early stage. Early detection, in turn, is vastly improved with community-wide organized sustainable programs. However, such schemes demand resources, and sizable amount of fund is being wasted for biopsy of benign lesions. Ultrasound imaging, because of its low ionization, effectiveness in diagnosing lesions of younger women and lesser resource intensiveness makes it a viable option for mass population programs. It has been shown to be effective in prevention of significant number of unnecessary biopsies. Thus, ultrasound imaging modality remains a prominent tool for diagnosis of breast cancer. Rigorous work is being done to improve overall imaging modality. In light of breast cancer, quantization of acoustic and morphometric features has proven to give good Receiver Operating Characteristic (ROC) area performance. However, ambiguity may still be there at concluding a lesion to be benign. The natural trend for such scenarios is the follow-up checkups recommended by experts. With thousands of follow-up checkups occurring each year, the need for a systematic study is imminent. We develop an effective algorithm that systematically processes over-time-data with the use of thin plate spline smoothing. This is followed by systematic categorization of different sets of physiological changes in light of benign and malignant lesions. Finally a versatile community wide scheme is outlined.

CHAPTER 1
INTRODUCTION

1.1 Global Concern of Breast cancer

Breast cancer is the top cancer in women worldwide and is increasing particularly in developing countries where the majority of cases are diagnosed in late stages. The incidence of breast cancer is increasing mainly due to increased urbanization and adoption of western lifestyles [1]. It is estimated that 519,000 women died in 2004 due to breast cancer, and although breast cancer is thought to be a disease of the developed world, a majority (69%) of all breast cancer deaths occurs in developing countries [2]. Breast cancer survival rates vary greatly world-wide, ranging from 80% or over in North America, Sweden and Japan to around 60% in middle-income countries and below 40% in low-income countries [3].

Although measures can be taken for prevention, majority of breast cancer cannot be eliminated. However, if breast cancer can be detected early, while it is still localized and before it can be palpated, the prognosis for cure is excellent [4]. Thus early detection is the key to reduce mortality rates.

With ambiguity in symptoms with other diseases and with lack of medical awareness, routine diagnosis is an important early detection strategy. This becomes even more important in low-income countries where resources to treat cancer patients at late stage are even more limited. There is some evidence that this strategy can produce "down staging" (increasing in proportion of breast cancers detected at an early stage) of the disease to stages that are more amenable to curative treatment [5].

1.2 Diagnostic systems

Mammography is the primary tool for routine diagnosis of breast cancer. Any lesion found is then further diagnosed by biopsy/ultrasound imaging or some other diagnostic tools. The concerns of prime importance are saving lives and saving resources, which is also indirectly correlated to saving lives.

Needle-point Biopsy is a lot more resource intensive than ultrasound imaging methods. However, current ultrasound imaging methods may yet falsely label a malignant lesion to be benign, although that possibility is significantly small. To alleviate these problems, experts recommend follow-up checkups and follow-up ultrasound imaging.

1.3 Ultrasound Imaging

Ultrasound is the second most widely used diagnostic imaging modality after x-ray. This imaging technology utilizes ultrasound to map the human organs. Medical physicists have discovered how to turn ultrasound imaging into a powerful tool for studying body structure and function. For example, medical ultrasound imaging is an important imaging method for obstetrics (imaging the fetus in the uterus during pregnancy), internal medicine (abdominal imaging, among others), cardiology (imaging the heart and circulatory system), and cancer imaging (distinguishing tumors from fluid-filled cysts in the breast and abdomen).

1.3.1 Working Principle of Ultrasound Imaging

Ultrasound imaging is based on the same principles involved in the sonar used by bats, ships and fishermen. When a sound wave strikes an object, it bounces back or echoes. Using a piezoelectric transducer a sound wave ranging from 1-10 MHz is produced. These pressure waves travel through the tissues and bones and reflect back from the boundaries, causing echo. The echo is then caught by the transducer and processed by the signal processing unit for imaging and display. To focus the sound waves towards a particular focal point, a set of transducer elements are energized with a set of time-delayed pulses to produce a set of sound waves that propagate through the region of interest, which is typically the desired organ and surrounding tissue.

Once the transducers have generated their respective sound waves, they become sensors that detect any reflected waves that are created when the transmitted sound waves encounter a change in tissue density within the region of interest. By properly time delaying the pulses to each active transducer, the resulting time-delayed sound waves meet at the desired focal point that resides at a pre-computed

depth along a known scan line. The amplitude of the reflected sound waves forms the basis for the ultrasound image at this focal point location.

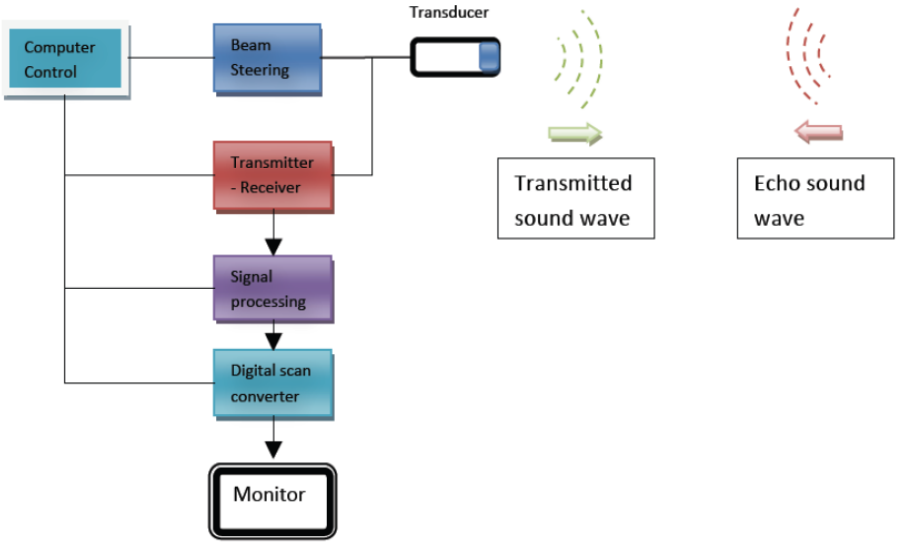


Fig. 1.1: Ultrasound working principle

1.3.2 Ultrasound Scanning Modes

There are typically six types of scanning modes available. These are:

1. A-MODE: A-mode (Amplitude) imaging displays the amplitude of a sampled signal for a single sound wave as a function of time. This mode is considered 1D and used to measure the distance between two objects by dividing the speed of sound by half of the measured time between the peaks in the A-mode plot, which represents the two objects in question. This mode is no longer used in ultrasound systems.

2. B-MODE: B-mode (Brightness) imaging is the same as A-mode, except that brightness is used to represent the amplitude of the sampled signal. B mode imaging is performed by sweeping the transmitted sound wave over the plane to produce a 2D image. Typically, multiple sets of pulses are generated to produce

sound waves for each scan line, each sets of pulses are intended for a unique focal point along the scan line.

3. M-MODE: An M-mode display refers to scanning a single line in the object and then displaying the resulting amplitudes successively. This shows the movement of a structure such as a heart. Because of its high pulse frequency (up to 1000 pulses per second), this is useful in assessing rates and motion and is still used extensively in cardiac and fetal cardiac imaging.

4. DOPPLER MODE: This mode makes use of the Doppler effect in measuring and visualizing blood flow.

•Color Doppler: Velocity information is presented as a color coded overlay on top of a B-mode image.

•Continuous Doppler: Doppler information is sampled along a line through the body, and all velocities detected at each time point is presented (on a time line)

•Pulsed wave (PW) Doppler: Doppler information is sampled from only a small sample volume (defined in 2D image), and presented on a timeline

•Duplex: a common name for the simultaneous presentation of 2D and (usually) PW Doppler information. (Using modern ultrasound machines color Doppler is almost always also used, hence the alternative name Triplex).

5. PULSE INVERSION MODE: In this mode two successive pulses with opposite sign are emitted and then subtracted from each other. This implies that any linearly responding constituent will disappear while gases with non-linear compressibility stand out.

6. HARMONIC MODE: In this mode a deep penetrating fundamental frequency is emitted into the body and a harmonic overtone is detected. In this way depth penetration can be gained with improved lateral resolution.

1.4 Biomedical Applications of Ultrasound

Ultrasound imaging is used in many types of examinations and procedures. Some examples include:

•Ultrasounds are useful in the detection of pelvic abnormalities and can involve techniques known as abdominal (transabdominal) ultrasound, vaginal (transvaginal or endovaginal) ultrasound in women, and also rectal (transrectal) ultrasound in men.

•Doppler ultrasound (to visualize blood flow through a blood vessel).

•Bone sonography (to diagnose osteoporosis).

•Ultrasound can also be used for elastography. This can be useful in medical diagnoses, as elasticity can discern healthy from unhealthy tissue for specific organs/growths.

•Echocardiogram (to view the heart).

•Fetal ultrasound (to view the fetus in pregnancy).

•Ultrasound-guided biopsies.

•Doppler fetal heart rate monitors (to listen to the fetal heart beat).

1.5 Advantages and Limitations of Ultrasound Imaging

As with all imaging modalities, ultrasound has its list of positive and negative attributes, which are listed below:

1.5.1 Advantages

- It is non-ionizing radiation, so it does not have the same risks as x-rays or other types of ionizing radiation.

- It images muscle, soft tissue, and bone surfaces very well and is particularly

useful for delineating the interfaces between solid and fluid-filled spaces.

- It renders "live" images, where the operator can dynamically select the most useful section for diagnosing and documenting changes, often enabling rapid diagnoses. Live images also allow for ultrasound-guided biopsies or injections, which can be cumbersome with other imaging modalities.

- It shows the structure of organs.

- It has no known long-term side effects and rarely causes any discomfort to the patient.

- Equipment is widely available and comparatively flexible.

- Small, easily carried scanners are available; examinations can be performed at the bedside.

- Relatively inexpensive compared to other modes of investigation, such as computed X-ray tomography, DEXA or magnetic resonance imaging.

- Spatial resolution is better in high frequency ultrasound transducers than it is in most other imaging modalities.

Through the use of an Ultrasound research interface, an ultrasound device can offer a relatively inexpensive, real-time, and flexible method for capturing data required for special research purposes for tissue characterization and development of new image processing techniques.

1.5.2 Limitations

•Ultrasound imaging devices have trouble penetrating bone. For example, imaging of the adult brain is very limited though improvements are being made in transcranial ultrasonography.

•Ultrasound imaging performs very poorly when there is a gas between the transducer and the organ of interest, due to the extreme differences in acoustic impedance. For example, overlying gas in the gastrointestinal tract often makes ultrasound scanning of the pancreas difficult, and lung imaging is not possible (apart from demarcating pleural effusions).

•Even in the absence of bone or air, the depth penetration of ultrasound may be limited depending on the frequency of imaging. Consequently, there might be difficulties imaging structures deep in the body, especially in obese patients.

•Body habitus has a large influence on image quality, image quality and accuracy of diagnosis is limited with obese patients, overlying subcutaneous fat attenuates the sound beam and a lower frequency transducer is required (with lower resolution)

•The method is operator-dependent. A high level of skill and experience is needed to acquire good-quality images and make accurate diagnoses.

•There is no scout image as there is with CT and MRI. Once an image has been acquired there is no exact way to tell which part of the body was imaged.

1.6 Background and Present State of the Problem

Breast cancer is the most common cancer among women other than non-melanoma skin cancer. Although survival rates for advanced-stage breast cancers have improved significantly, early-stage breast cancers are now virtually curable [6]. Thus increased awareness, early detection and cheaper diagnosis are the most

important factors in fighting against this disease. Ultrasound imaging, with comparatively lower cost equipments, non-invasive diagnosis and near accurate separation of benign lesions has a big potential for economic, comfortable and fast diagnosis of breast lesions.

Currently, in the field of medical diagnosis, ultrasound is responsible for about one in five of all diagnostic images [7]. Ultrasound elastography is emerging with enormous potential as a medical imaging tool for effective discrimination of pathological changes in soft tissue. It maps the tissue elasticity or strain due to a mechanical de-formation applied to it. Elastography by now is well recognized as an effective means for discriminating biological tissue pathology change in medical diagnosis. Because of high correlation of tissue stiffness with its pathology change, this new noninvasive imaging modality is showing a great promise in the detection and/or characterization of breast and prostate tumors [8]. Calibrated spectrum analysis procedures show the potential of providing key tissue descriptors that can complement high-quality B-mode images. Analysis has been done that provides a theoretical basis for interpreting spectral parameters in terms of physical attributes of tissue morphology [9].

To improve the accuracy of breast ultrasound diagnosis, American College of Radiology (ACR) developed the Breast Imaging Reporting and Data System (BI-RADS) lexicon of features describing the ultrasound appearance of breast lesions [10, 11 and 12]. Research is being done to develop methods to make breast-lesion evaluation quantitative, reproducible and relatively operator independent [10]. Several groups have studied automated methods of breast-cancer identification. For example, artificial neural-network-based classifier that uses several shape features such as branch patterns and number of lobulations and reported a Receiver Operating Characteristic (ROC)-curve area of 95% [13]. Multi-feature analysis based on BI-RADS criterion shows promise with its high ROC-curve area along with flexibility of adjusting ‘number of biopsy-falsely diagnosed benign lesion’ tradeoff by means of different thresholding [10].

However, all methods currently available still may falsely conclude a malignant lesion to be benign, although that probability is very low. Thus experts often recommend follow-up checkups before concluding a lesion to be benign.

1.7 Motivation

Independent of diagnostic procedures available, the most important key to the success of early detection are careful planning and a sustainable and well organized coordination across the whole community [1]. A cheaper procedure which is less resource and manpower intensive, as in non-invasive diagnosis, is actually the answer for community-wide sustainable programs.

However, all methods currently available may still falsely conclude a malignant lesion to be benign, although that probability is very low. We propose quantification of follow-up checkups, already in use among doctors, to alleviate that problem. This procedure, if implemented, should drastically improve robustness of isolating significant ratio of benign lesions, and the robustness is statistically expected to improve with increasing database. Furthermore, this procedure has the potential to isolate different types of physiology and lesion behavior. This, in particular may provide insight into disease mechanisms and help formulate new medicines for treatment/prevention. This algorithm that we have developed may also be applied to numerous other applications for relating different statistical phenomenon.

Finally this method is another step forward to automated breast cancer diagnosis, with a potential to save costs incurred from thousands of medical fee paid to interpret results from expert pathologists for every follow-up checkup.

1.8 Outline of the Dissertation

- Chapter – 1 gives a brief idea about Breast Cancer and application of Ultrasound Medical Imaging Modality for early isolation of benign lesions.
- Chapter – 2 discusses acoustic and morphometric features of benign against malignant breast cancer.
- Chapter – 3 illustrates the method of quantification of acoustic and morphometric features.
- Chapter – 4 outlines a novel method for more robust isolation of benign from malignant and shows simulation results.
- Chapter – 5 demonstrates the simulation and findings of the thesis.
- Chapter – 6 discusses scope and future work**.**

CHAPTER 2

BENIGN AGAINST MALIGNANT LESION

2.1 Overview

A benign tumor is a well-defined growth with smooth boundaries. This type of tumor simply grows in diameter. A malignant tumor usually has irregular boundaries and invades the surrounding tissue. An illustration is show in figure 2.1.

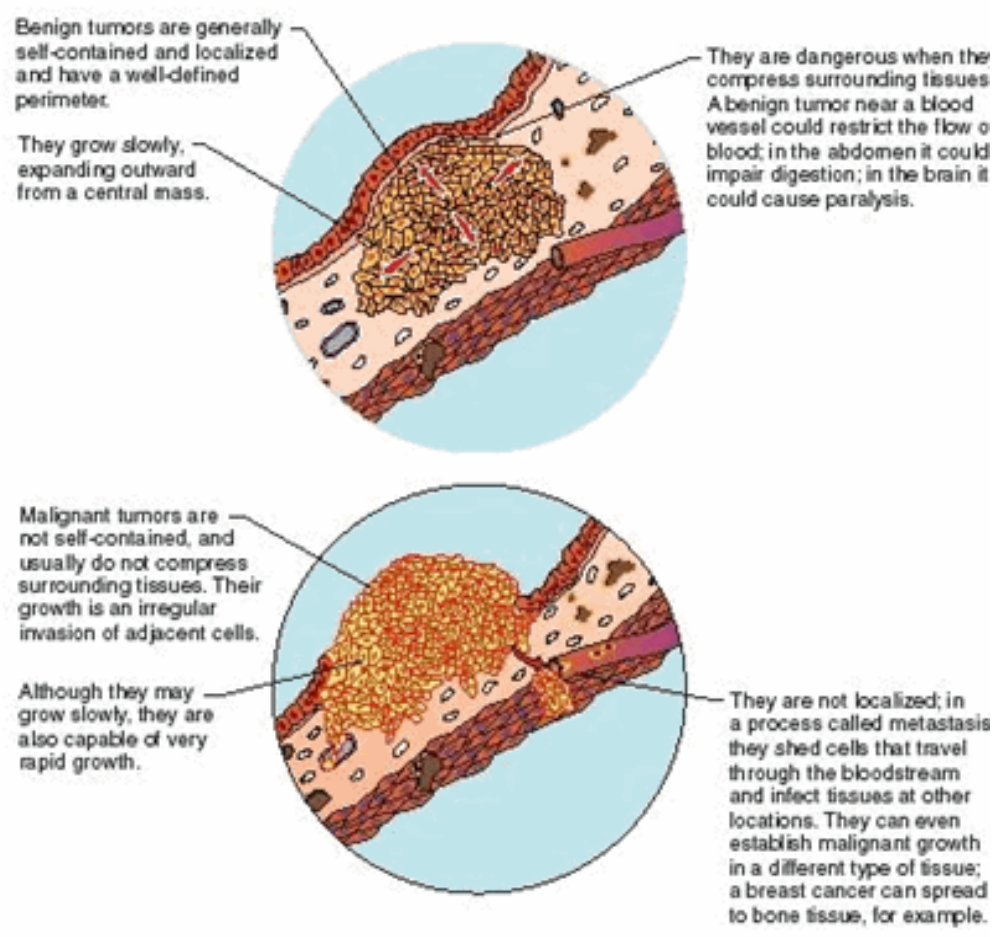


Fig. 2.1: Benign tumor versus Malignant Tumor [14]

2.2 Illustration of the Lexicon

The Breast Imaging Reporting and Data System: Ultrasound was developed by the Breast Ultrasound Lexicon Subcommittee of an Expert Working Group, with support from the Office of Women's Health, National Institutes of Health and the Commission on Ultrasound of the American College of Radiology. Techniques adapted from those used in the development of BI-RADS™ were used to develop the Breast Imaging Reporting and Data System: Ultrasound [15]. This lexicon is intended to assist users in applying standardized descriptors to characterize and report sonographic features of breast masses. The following articles talk about the different characterization of breast lesions alongside illustrations of figures 2.2-2.15 [15].

2.2.1 Mass shape

Mass shape is characterized in terms of round, oval, or none of these. Benign lesions, for example simple cysts, are typical of spherical shaped lesions. For malignant lesions and fibroadenomas, the shape of the masses cannot be classified as either round or oval. Unlike typical lumps from breast cancer, fibroadenomas are easy to move, with clearly defined edges.

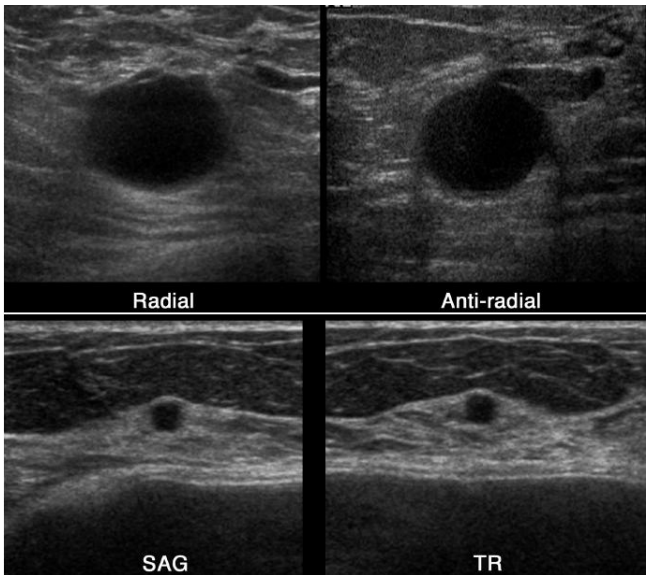


Fig. 2.2 Large and small simple cysts, each with a typical spherical shape

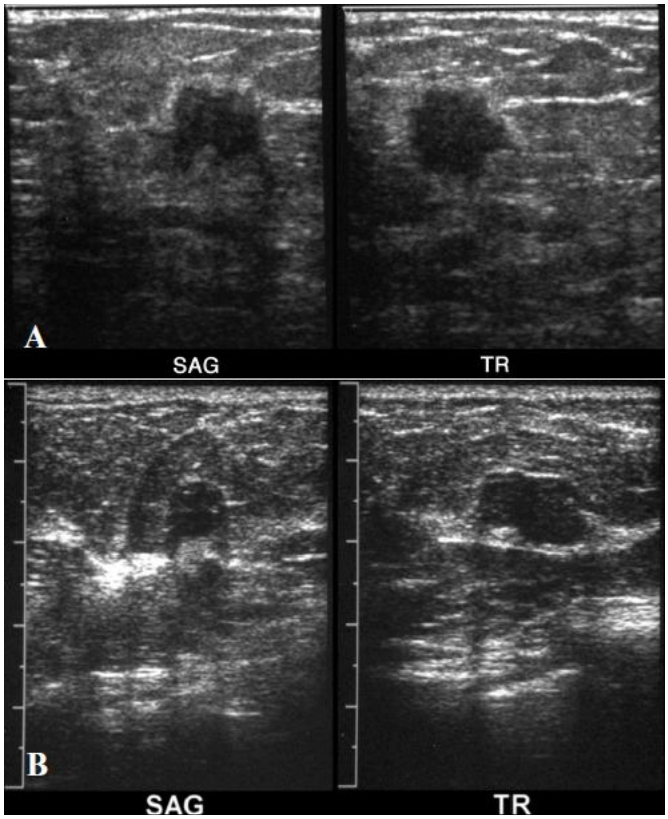


Fig. 2.3 (A) Invasive carcinoma (B) fibroadenoma.

2.2.2 Mass Orientation

Mass orientation is classified in terms of parallel characterization of longer axis of lesion to the surface.

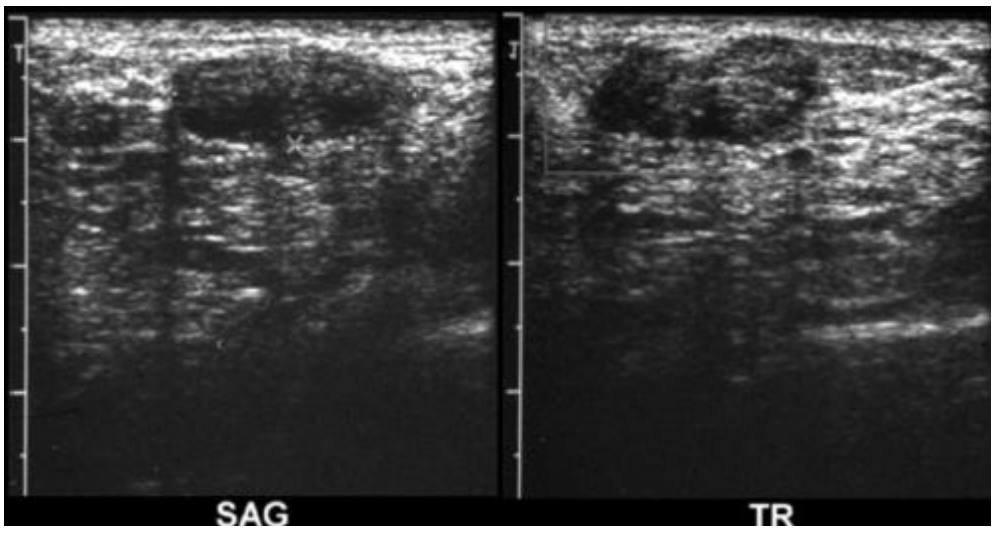


Fig. 2.4: Fibroadenoma. The long axis of the mass is parallel to the surface

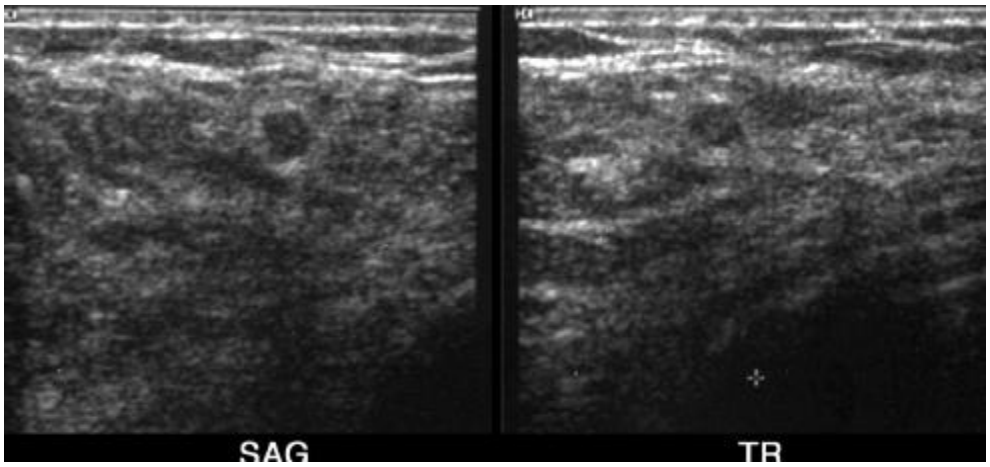


Fig. 2.5: Fibroadenoma. The long axis is not parallel.

2.2.3 Mass Margins

The margin is the boundary between the lesion and its surroundings. Malignant lesions are generally characterized by indistinct or ill-defined margins.

1. Distinct (Circumscribed)

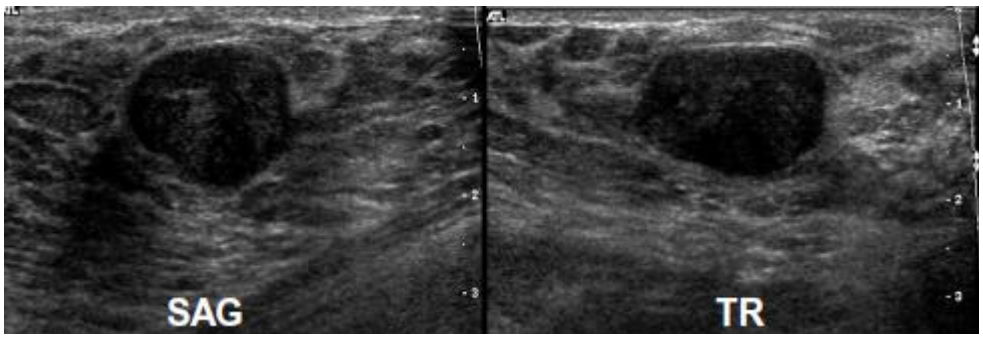


Fig. 2.6: Fibroadenoma with distinct circumference.

B. Indistinct

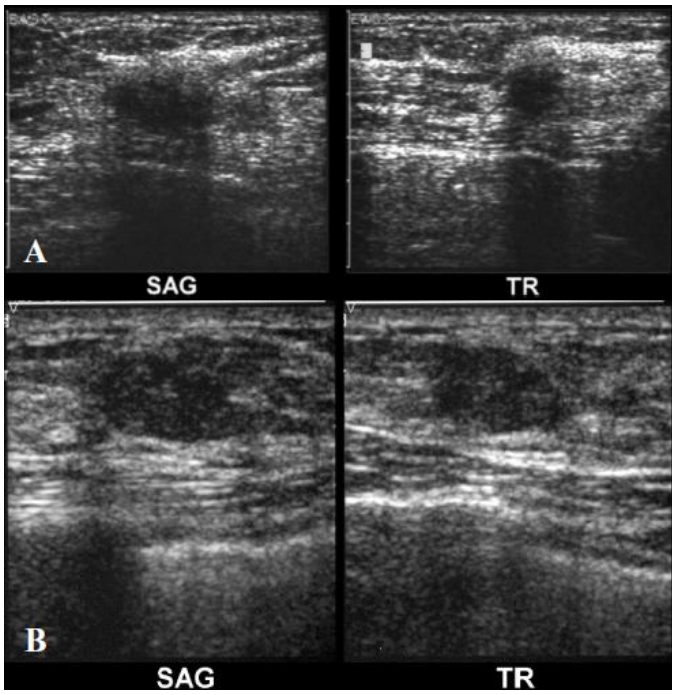


Fig. 2.7: Suspicious lesion with indistinct circumference.

In figure 2.7(A) the interface of the mass with the surrounding tissue is ill defined. In figure 2.7(B) portions of the margin of the mass are defined and other portions are indistinct. Since ill-defined margins carry a higher degree of suspicion for malignancy than well-circumscribed margins, this mass should be classified in accord with its most suspicious feature - indistinct margins.

2.2.4 Mass Echogenicity / Echotexture

A. Homogeneous

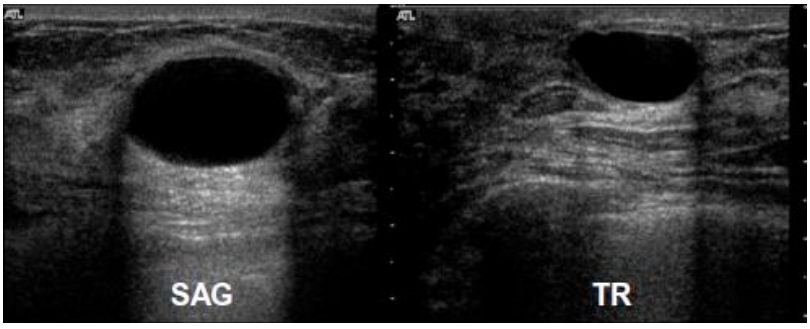


Fig. 2.8: Simple cyst

B. Heterogeneous

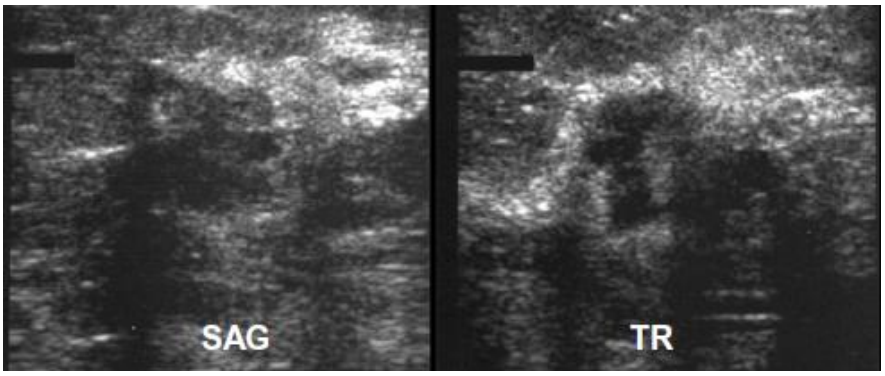


Fig. 2.9: Invasive cancer.

Although predominately hypoechoic, scattered areas if increased echogenicity are present in the mass.

C. Complex (Predominately cystic)

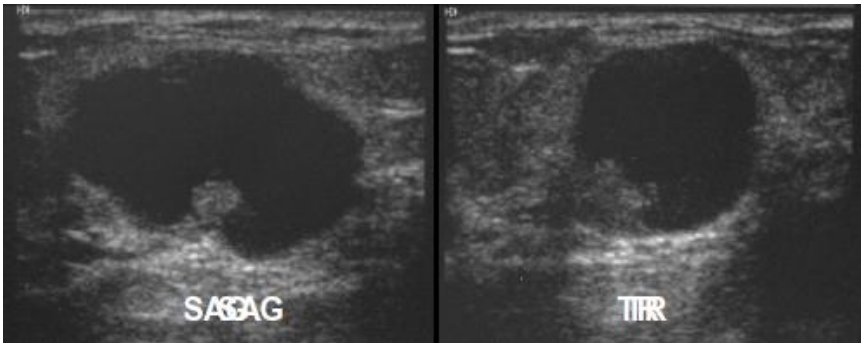


Fig. 2.10: Intracystic cancer.

Predominately cystic mass with additional feature of mural nodule is representing the cancer in here. Additional features of complex masses may be described if applicable include septations and mural nodules.

2.2.5 Shadowing/Enhancement

A. None

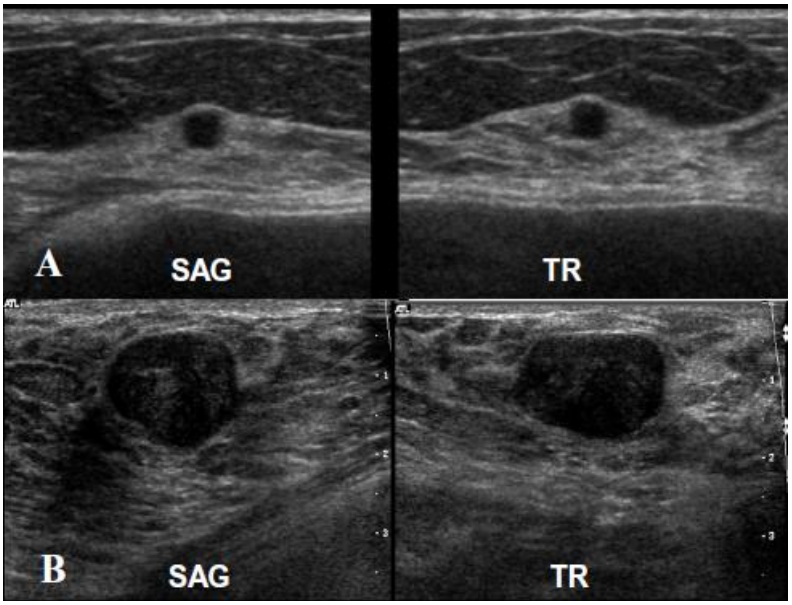


Fig. 2.11: (A) Small cyst and (B) fibroadenoma showing no evidence of significant attenuation or enhancement.

B. Enhancement

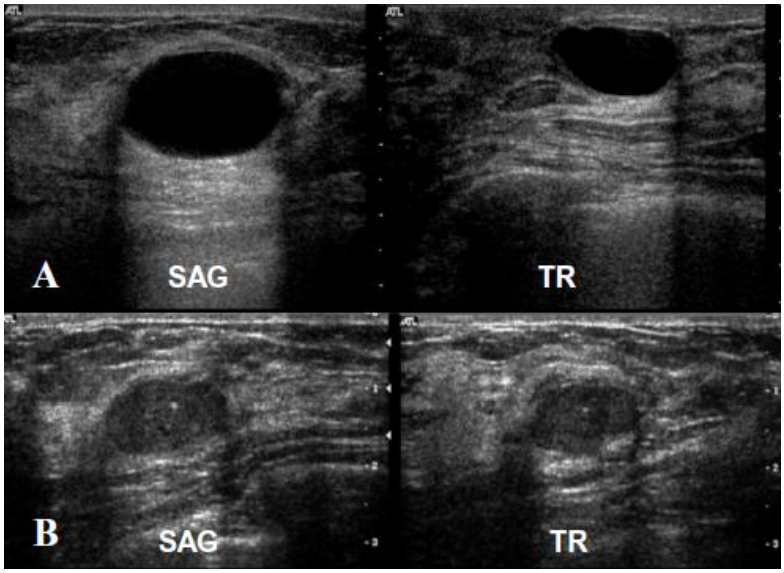


Fig. 2.12: (A) Large cyst (B)The fibroadenoma showing enhancement

C. Shadowing

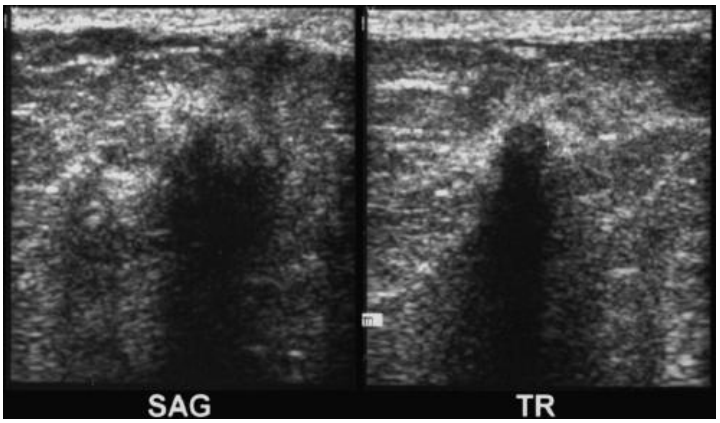


Fig. 2.13: Carcinoma with prominent central shadowing.

D. Combined Pattern (both shadowing and enhancement)

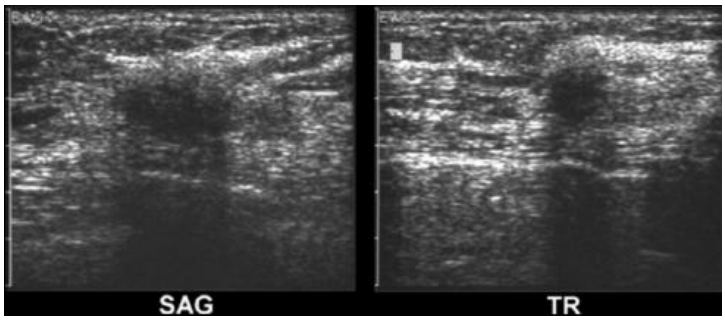


Fig. 2.14: Carcinoma showing a mixture of shadowing, no effect, and minimal enhancement.

2.2.6 Effect on Surrounding Tissue

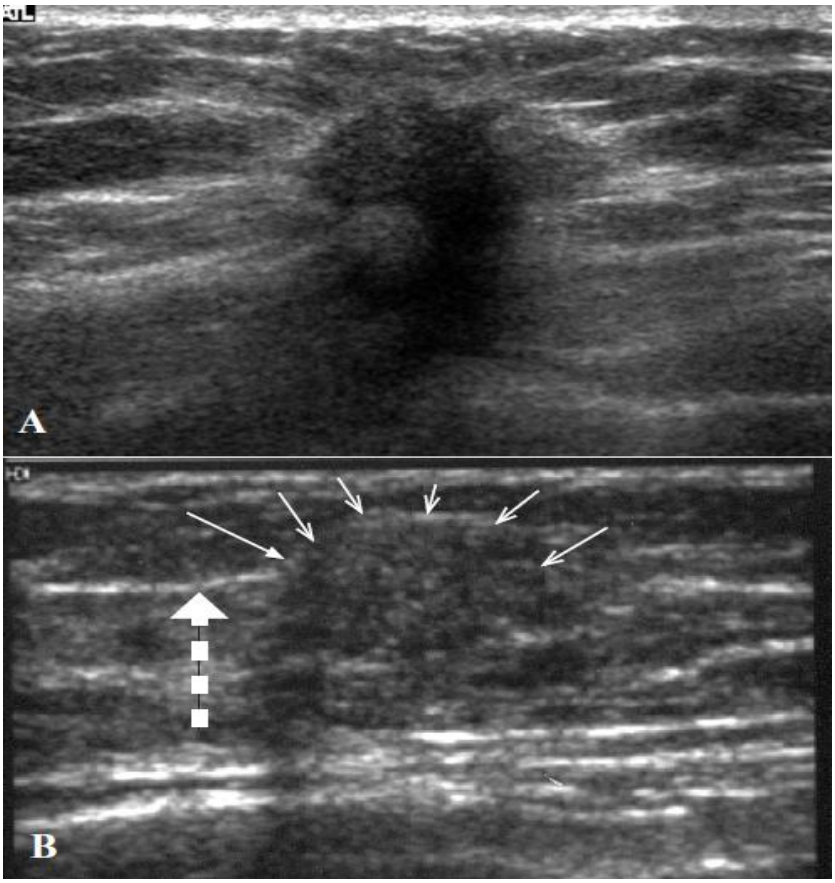


Fig. 2.15: Effect on surrounding tissue.

Invasive carcinoma in figure 2.15(A) distorts normal breast architecture byangulation, retraction, and thickening of Cooper ligaments. In figure 2.15(B) a subtle mass (solid arrows) produces abrupt interruption (broken arrow) of a Cooper ligament.

2.3 Summary and Implementation of multi-feature analysis

Summary of the acoustic and morphometric features in contrast between benign and malignant lesions are outlined in table 2.1. This features, if studied and efforts taken to quantize, can lead to robust isolation of significant amount of benign lesions. The following chapter illustrates exactly how a multi-feature analysis is applied in regard of this problem.

Table 2.1 Features of conventional B-Mode images that are typically associated with malignant and benign lesions (BI-RADS criteria) [10].

Malignant lesions			**Benign lesions**	
Morphometric features	***Acoustic features***		***Morphometric features***	***Acoustic features***
Irregular shape/spiculation	Central shadowing		Spherical/ovoid shape	Edge shadowing or enhancement
Poorly defined margin	Hypoechogenicity		Linear well-defined margin	Hyperechogenicity
Tall aspect ratio	Heterogeneous texture		Thin capsule	Homogeneous texture
Microlobulation	Calcifications		Gentle bi- or trilobulations	
Architectural Distortion			Orientation parallel to tissue plane	

CHAPTER 3
MULTI-FEATURE ANALYSIS OF SOLID BREAST LESION

3.1 Method

Quantitative descriptors have been developed to provide a primary objective means of identification of cancerous breast lesions. These descriptors include quantitative acoustic features and morphometric features. Acoustic features include echogenicity, heterogeneity and shadowing. These are computed by generating spectral-parameter images of the lesion and surrounding tissue. Morphometric features are quantized by geometric and fractal analysis of traced lesion boundaries [10].

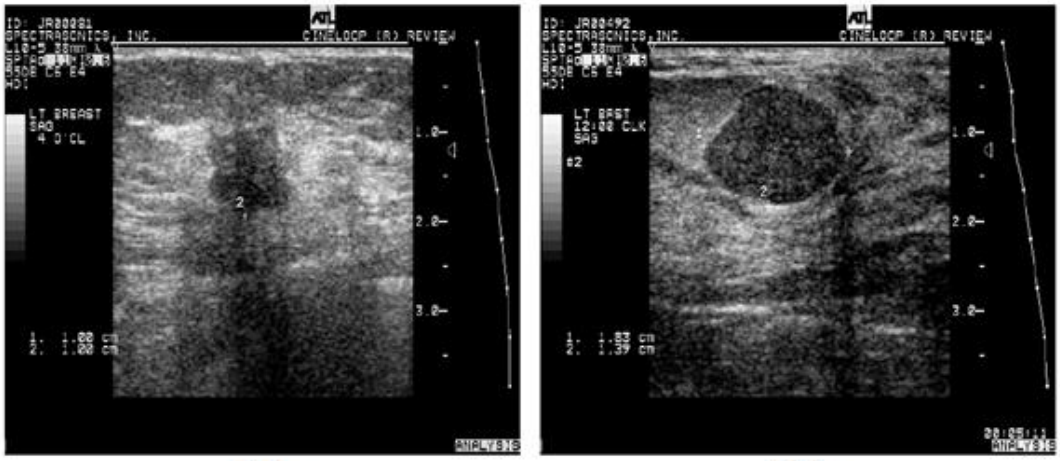


Fig. 3.1: (A) Malignant lesion, (B) Benign lesion (fibroadenoma).

In figure 3.1(A), the lesion has an irregular multilobular shape with a ‘tall’ aspect ratio (i.e., has a higher ratio of height to width), heterogeneous internal texture, and a prominent shadow. In figure 3.1(B), the lesion has the classical near-spherical shape, a smooth boundary, homogeneous internal texture and a posterior enhancement or ‘anti-shadow’. Figure 3.2 shows the segmentation and traces for further analysis. Tissue characterization will be next based on acoustic and morphometric features, defining tissue shape and behavior to high frequency sound, i.e. ultrasound.

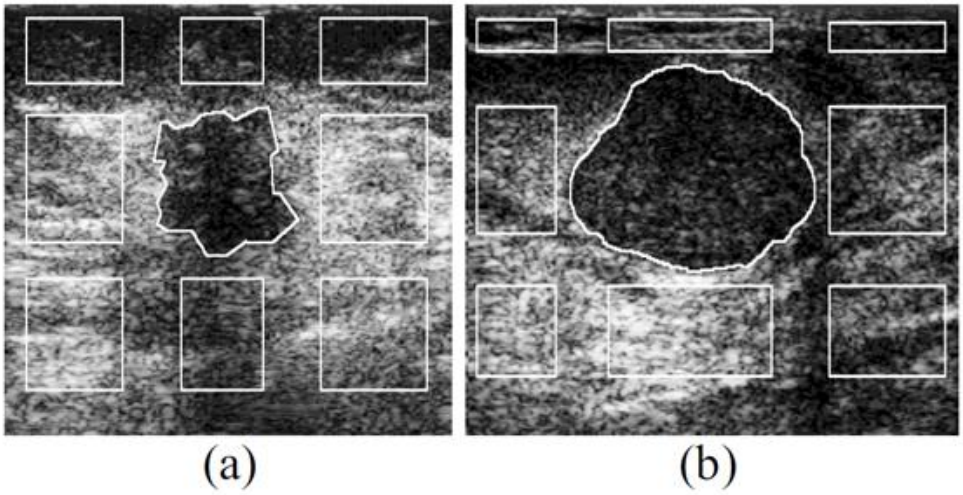


Fig. 3.2: B-scan images of figure 1 with analysis-region traces superimposed. (a) Malignant lesion. (b) Benign lesion.

3.2 Acoustic features

Acoustic features have been quantified in terms of calibrated spectrum-analysis. This involves several steps. Hamming window to rf data, followed by Fourier transform and finally expressing resultant power spectrum in dB. A linear regression line, best fit to the power spectrum, is then found. Mid-band power spectrum value, M, mid-band frequency, f_0, and spectral intercept I is then computed, as illustrated in figure 3.3. These parameters are used in algorithms to produce different acoustic features.

Echogenicity, or ability to bounce an echo, is quantified by the mean spectral intercept, i.e. mean of I. This relates to the fact that higher fraction of echo from a tissue point at a particular window relates to an overall higher average of mean spectral intercept of power spectrum. Heterogeneity of texture in a lesion has been quantified by use of the values of M, mid-band power spectrum, from all the windows taken inside a lesion. Next four-neighborhood pixel algorithm (FNPA) is chosen instead of histogram algorithm (HA) since FNPA can reflect all textural differences [16]. For each window FNPA makes use of values from the four

adjacent windows to its edges to yield FP1/μ, where FP1 is average of differences between M values and μ is the mean of FP1 across the entire lesion.

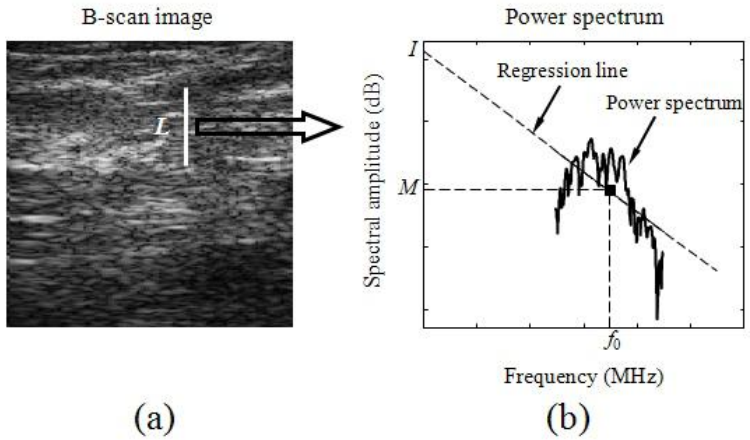


Fig. 3.3: Illustration of spectrum analysis procedure. (a) B-Scan (breast) image of a patient. (b) Power spectrum

3.3 Morphometric feature

All morphometric features have been computed using lesion boundaries traced on B-mode images. Aspect ratio has been defined as the maximum vertical lesion dimension divided by maximum horizontal lesion dimension (depth divided by width). Fuzziness of borders is quantized by margin definition which is defined as the sum of the magnitude of the gradient of M on a lesion contour normalized by the sum of magnitude of M on the lesion contour.

Morphometric features related to the shape or border of the lesion [17] demonstrated good performance (97.2% sensitivity and 94.1% negative predictive value) using quantitative lesion-shape features describing irregular boundaries, quantified here using fractal dimension ratio. Fractal dimension ratio, which represents the amount of complexity in a pattern, can be used as a measure of border irregularity, since irregular border means more complexity. This fractal dimension ratio is calculated using Hurst Coefficient since $HD = 2 - H$, where H is the Hurst exponent coefficient [18] and HD is the Fractal (Hausdorf) Dimension. The irregularity of borders is illustrated against Fractal (Hausdorf) Dimension, HD, and

convexity or ratio of border points with internal angles less than 180°, C, in figure 3.4.

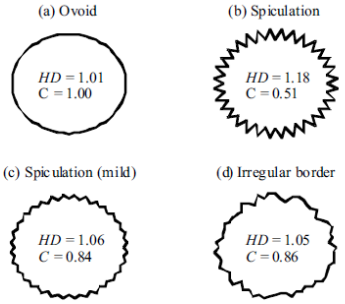


Fig. 3.4: Illustration of border irregularity and fractal (Hausdorf) dimension and convexity.

3.4 Case study of 130 biopsy-proven patients

As discussed in earlier section, rf echo-signal from ultrasonography is processed to quantify different acoustic and morphometric features. For each feature a null hypothesis is tested that the feature value does not differ for benign and malignant cases in the population. If the significance level is low then the null hypothesis can be rejected. The significance levels are low for FNPA, margin definition, aspect ratio and fractal (Hausdorf) dimension. Thus, we can use these features to isolate a significant ratio of benign cases by using a straight line divider achieving a ROC-curve area of 0.947+0.045 [10]. Figure 3.5 illustrates these feature values processed for 130 biopsy proven patients.

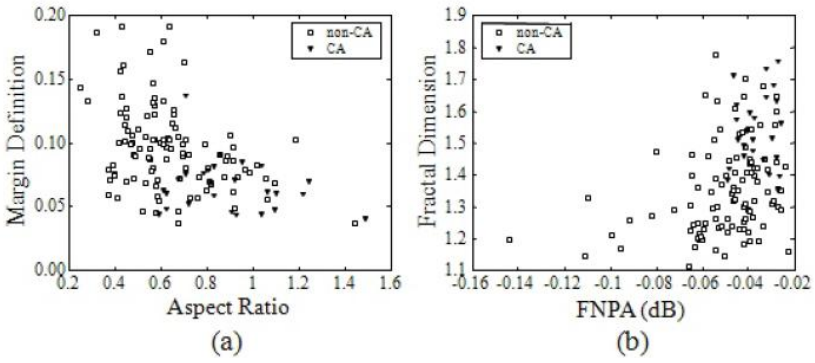


Fig. 3.5 Shows feature values of margin definition, aspect ratio, fractal dimension and FNPA of 130 biopsy-proven patients. Benign-proven patients are represented as empty-squares and malignant-proven patients are represented as blacked-triangles [10].

Several other groups have studied automated methods of breast-cancer identification. Shankar et al describes an approach based on the Nakagami Distribution [20] and non-Raleigh statistics of the envelope [21]. (An ROC area of 0.8701 ± 0.0345 is reported for the latter study.) Drukker et al [22] describes a computerized analysis of shadowing to classify breast lesions and reported a detection sensitivity of 80%. Joo et al [23] described an artificial, neural-network-based classifier that uses several shape features such as spiculation, branch patterns and number of lobulations and reported an ROC-curve area of 95%.

CHAPTER 4

A SYSTEMATIC STUDY OF FOLLOW-UP CHECKUPS

4.1 Methodology

In previous chapter, a straight line could be drawn that separates benign lesions only. However, doubts may arise when the scatter point of the new patient is in certain regions. Medical experts tend to recommend a follow-up checkup. This section of our work is to find a systematic study of that follow-up checkup and quantification of the results. We use interpolation techniques and linear correlation to our cause. The following articles and figure 4.1-4.3 illustrates the different interpolation techniques [24].

4.2 Interpolation techniques

For diagnosed data of a new patient, it is required to **interpolate** (i.e. estimate) the value of that function for an intermediate value of the independent variable. This is achieved by curve fitting or regression analysis [24].

4.2.1 Linear interpolation

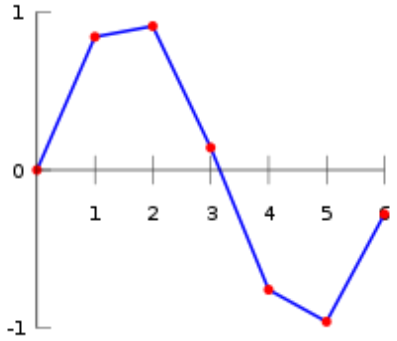


Fig. 4.1: Plot of the data with linear interpolation applied

Linear interpolation, as in figure 4.1, takes two data points, say (x_a,y_a) and (x_b,y_b), and the interpolant at the point (x,y) is given by:

$$y = y_a + \frac{(y_b - y_a)((x - x_a))}{(x_b - x_a)}$$

Linear interpolation is quick and easy, but it is not very precise. Another disadvantage is that the interpolant is not differentiable at the point x_k (i.e., sampling points).

4.2.2 Polynomial interpolation

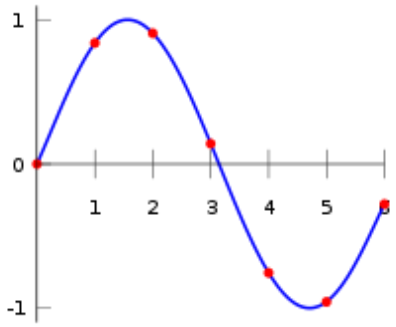


Fig. 4.2: Plot of the data with polynomial interpolation applied

Let us consider again the problem with data set in figure 4.2. For polynomial interpolation, the interpolation error is proportional to the distance between the data points to the power *n*. Furthermore, the interpolant is a polynomial and thus infinitely differentiable. So, we see that polynomial interpolation overcomes most of the problems of linear interpolation.

However, polynomial interpolation may exhibit oscillatory artifacts, especially at the end points (Runge's phenomenon). These disadvantages can be reduced by using spline interpolation or restricting attention to Chebyshev polynomials.

4.2.3 Spline interpolation

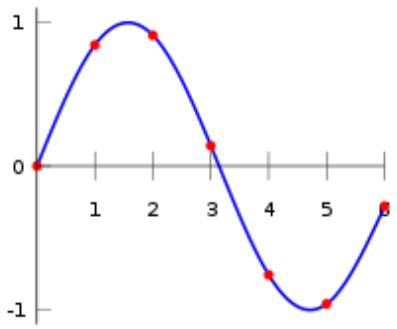


Figure 4.3: Plot of the data with spline interpolation applied

Spline interpolation is in theory a hybridization of linear interpolation and polynomial interpolation. As in linear interpolation which uses a linear function for each of intervals $[x_k,x_{k+1}]$, Spline interpolation uses low-degree polynomials in each of the intervals, and chooses the polynomial pieces such that they fit smoothly together. The resulting function is called a Spline.

The natural cubic spline interpolating the points for the same example considering the first four points only [24]:

$$f(x) = \begin{cases} -0.1522x^3 + 0.9937x, & if\ x \in [0,1] \\ -0.01258x^3 - 0.4189x^2, & if\ x \in [1,2] \\ 0.1403x^3 - 1.3359x^2 + 3.2467x - 1.3623, & if\ x \in [2,3] \end{cases}$$

Like polynomial interpolation, spline interpolation exhibits a smaller error than linear interpolation and the interpolant is smoother and it also does not suffer from Runge's phenomenon. For our multivariable case, among the different types of spline techniques available, we opted for thin plate spine particularly because the model does not need manual tuning of parameters. This makes it appropriate for an automated quantification of follow-up checkups.

Thin plate spline (TPS) fits a mapping function *f(x)* between corresponding point-sets $\{y_i\}$ and $\{x_i\}$ by minimizing the following energy function [25]:

$$E = \iint \left[\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial xy} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right] \mathrm{d}x \, \mathrm{d}y$$

And for a smoothing TPS, it is:

$$E_{tps} = \sum_{i=1}^{K} \| y_i - f(x_i) \|^2 + \lambda \iint \left[\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right] \mathrm{d}x \, \mathrm{d}y$$

Then smoothing TPS is defined as: $f_{tps} = arg \{ min (E_{tps}) \}$

The name thin plate spline refers to a physical analogy involving the bending of a thin sheet of metal. In the physical setting, the deflection is in the *z* direction, orthogonal to the plane. In order to apply this idea to the problem of coordinate transformation, one interprets the lifting of the plate as a displacement of the *x* or *y* coordinates within the plane. This is related to this thesis by using the *z* direction to be considered to be time axis, and the lifting of the plate to be related to the different feature value displacements in successive follow-ups.

4.2.4 Splines in Matlab

In the simplest situation, one is given points and is looking for a piecewise-polynomial function that satisfies all, more or less. An approximate fit might involve least-square approximations or the smoothing spline. For our cause we use a multivariate function, the stform, which uses arbitrary or scattered translates $\Psi(x\text{-}c_j)$ of one fixed function Ψ, in addition to some polynomial terms [26].

$$f(x) = \sum_{j=1}^{n-k} (\Psi(x - c_j) a_j + p(x))$$

Explicitly, such a form describes a function in terms of the basis function Ψ, a sequence (c_j) of sites called centers and a corresponding sequence of coefficients

(a_j) of n coefficients, with the final coefficients, $a_{jn-k+1},\ldots a_n$ involved in the polynomial part, p.The most important property of stform is the function f depends linearly on its coefficients.

$$f(x) = \sum_{j=1}^{n} (\Psi_j(x) a_j)$$

Here Ψ_j is either translates of the basis function Ψ or else is some polynomial.

We use 'tpaps' stform which enables us to create thin-plate spline approximations, f, that satisfy, approximately or exactly, the equation $z=f(x,y)$ for given data values, z, at given scattered data sites (x,y) in the plane.

4.3 Correlation

To relate between the different best-fitted data, a Pearson's correlation coefficient is actually a simple yet elegantly effective tool. The reason behind its effectiveness is the linear dependence of the best-fit stform with its coefficients.

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

In our procedure, we isolate cases that are strongly correlated with each other. This will result in different sets of malignant lesion with different physiological characteristic changes. A similar procedure is applied to benign lesions having strong correlation.

CHAPTER 5

SIMULATION AND INTERPRETATION OF THE FINDINGS

5.1 Studying some actual cases

We analyze two likely images shown in figure 5.1 that exhibit characteristic changes from a doubtful lesion to a benign looking lesion.

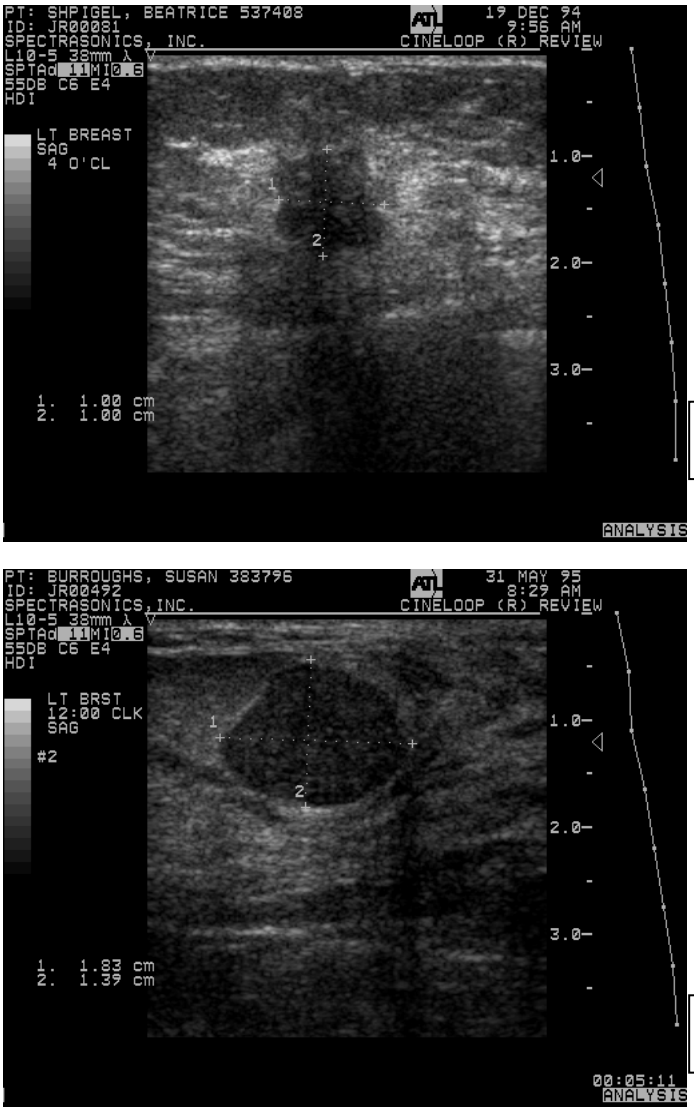


Fig. 5.1: One plausible case of a doubtful lesion over time (which later turns out to be benign)

Table 5.1: Corresponding quantification of images of figure 5.1

Quantified feature	***Image A***	***Image B***	***Change from Image A***
Aspect ratio	0.95	0.75	Decrease by 0.2
Margin definition	0.145	0.215	Increase by 0.07
FNPA(dB)	-0.06	-0.07	Decrease by 0.01
Fractal dimension	1.28	1.25	Decrease by 0.03

The second image exhibits lower aspect ratio, better defined border (lower value of margin definition), less heterogeneous texture inside outlined area (lower value of FNPA), and less irregular border (lower value of Fractal Dimension). If the second image was taken from same patient after an interval of time, it is obvious that the lesion is exhibiting characteristics of benign lesion. Although, this decision can be made by any expert pathologist, considering thousands of follow-up checkups every year, a quantification and subsequent automation of decision making can save resource and time. Even more importantly, changes may be a lot more subtle and making any concluding remark of tissue characterization can be difficult.

5.2 Need of perceptive simulation

Unlike cases illustrated in previous article, over-time images may not always exhibit such conclusive changes. This creates ambiguity over the decision for next step of diagnosis/treatment. Thus we aim for a robust modeling method for each patient and simultaneous correlation to different sets of benign/malignant cases. However, real data over time for individual patients has not yet been collected. Thus we build a fail-safe algorithm that becomes smarter with improved database.

5.3 Creating phantom data

The region of our interest is where there remains too much ambiguity for conclusive decision and does not seem to follow any regular pattern for either malignant or benign lesions as shown in figure 5.3. We, therefore, accept this limitation and create our phantom data in random manner. 60 biopsy-proven patients, either malignant-proven or benign-proven, exhibiting 3 types of over-time characteristic changes are simulated. Three new patients exhibiting similar changes over-time is then simulated and an automated classification applied by use of previous data of 60 biopsy-proven patients. Simulated feature values for the first ultrasonography diagnosis, i.e. first check-up, of all patients is shown in figure 5.3, and feature values for all the over-time follow up scans of one patient is shown in figure 5.4.

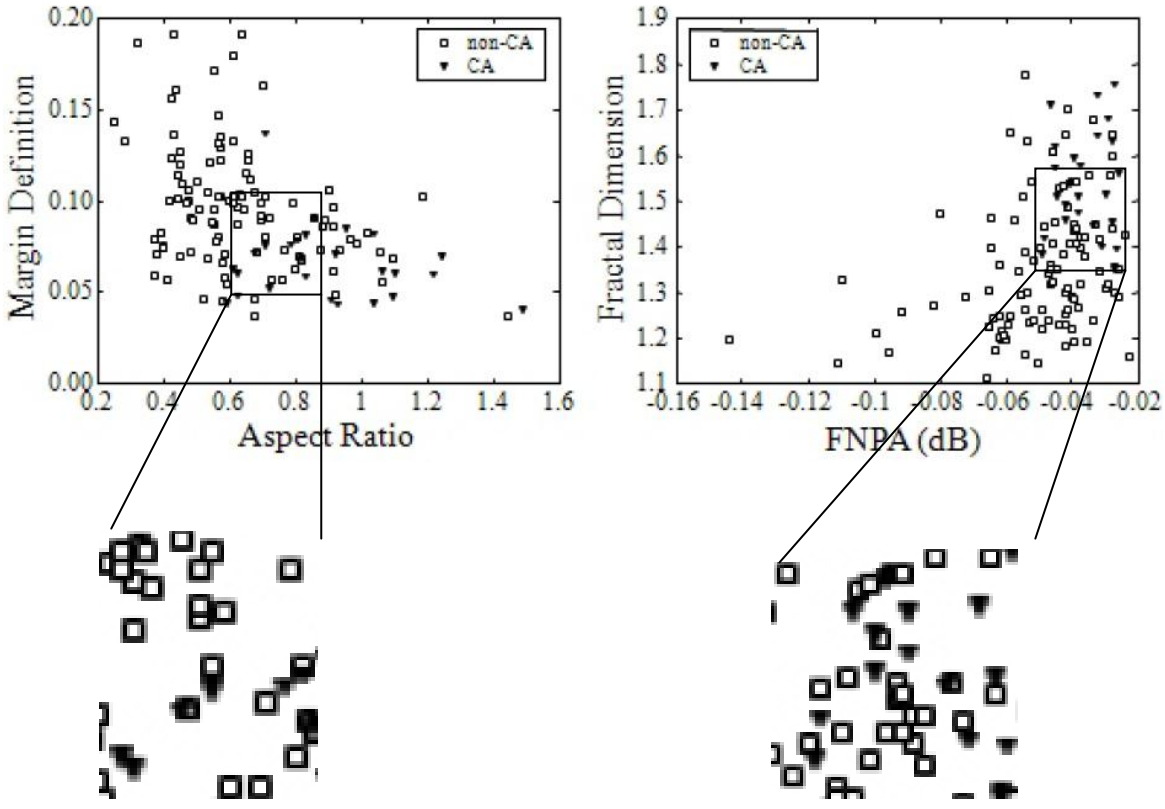


Fig. 5.2: Some of the ambiguous regions of scatter diagrams.

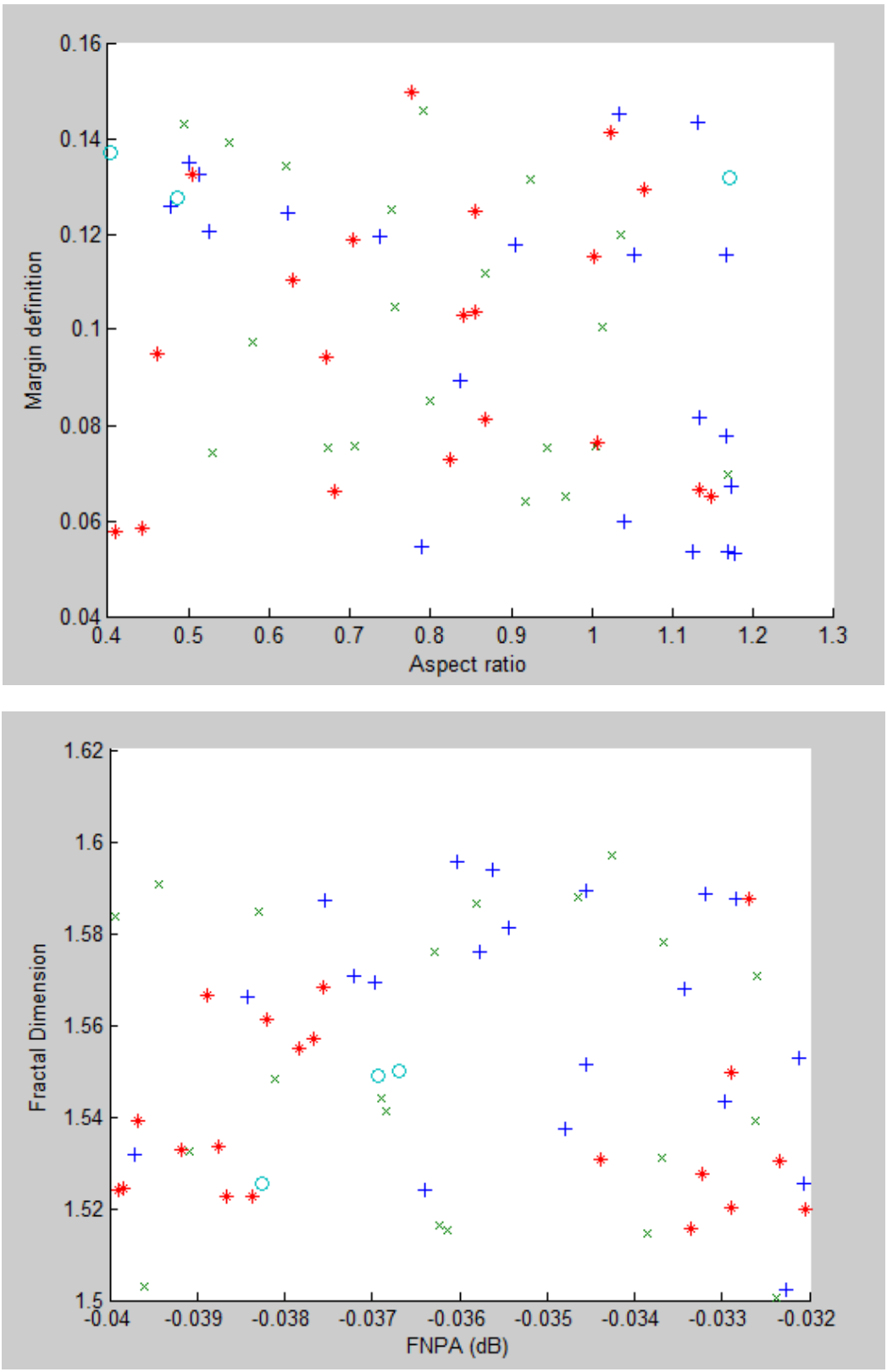


Fig. 5.3: Sixty-three phantom data generated (o-new patient, + is benign type A, * is benign type B, x is malignant type A)

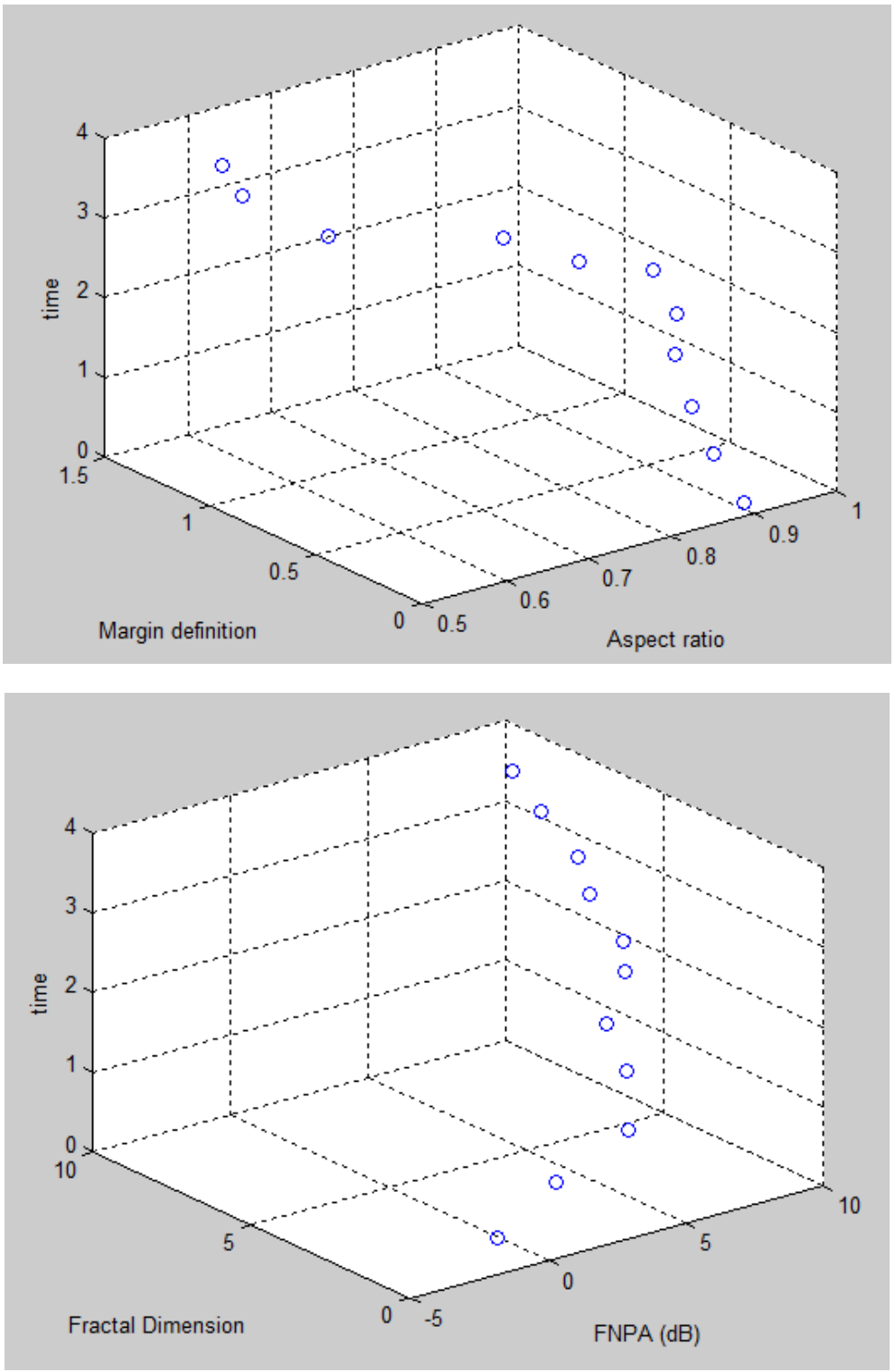


Fig. 5.4: Phantom time-spaced data generated for particular patient exhibiting non-linear change

5.4 Best-fit Modeling

The next step is to approximate the physiological changes of the patient illustrated in figure 5.4 into a best-fit.

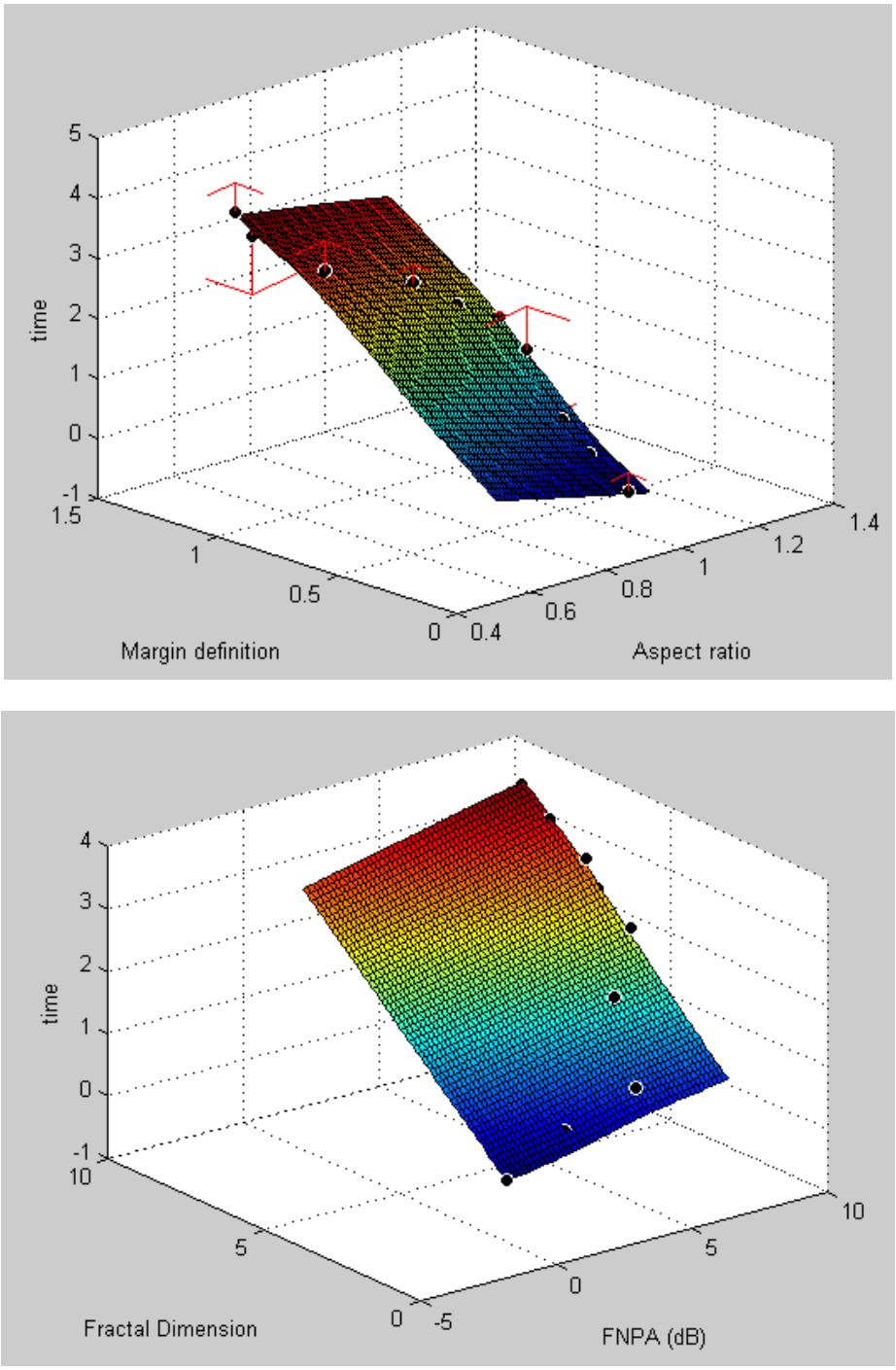
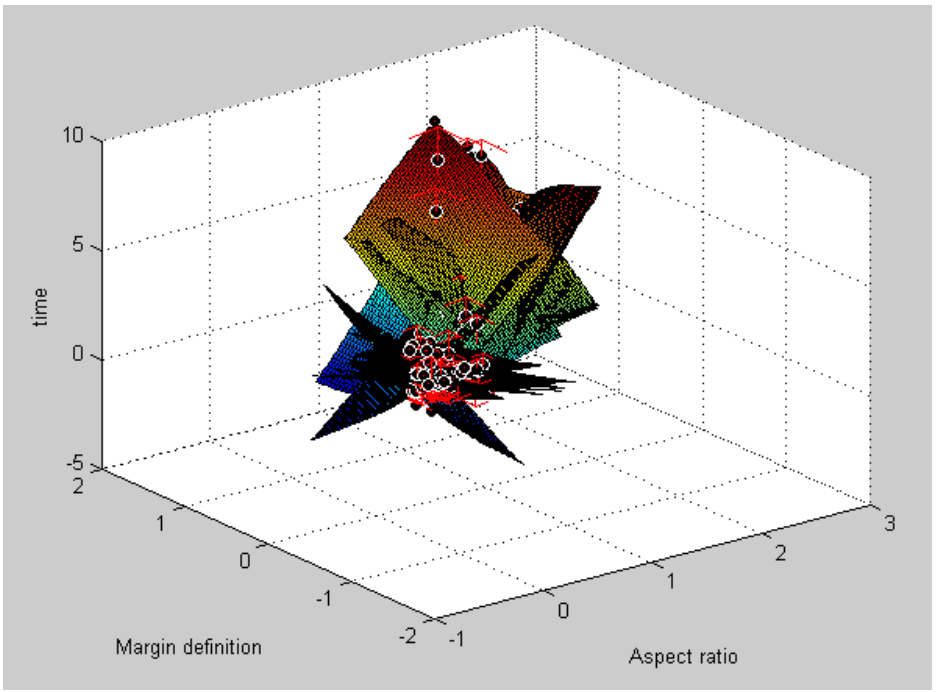


Fig. 5.5: Best-fit stform or modeled pattern for the patient of figure 5.4



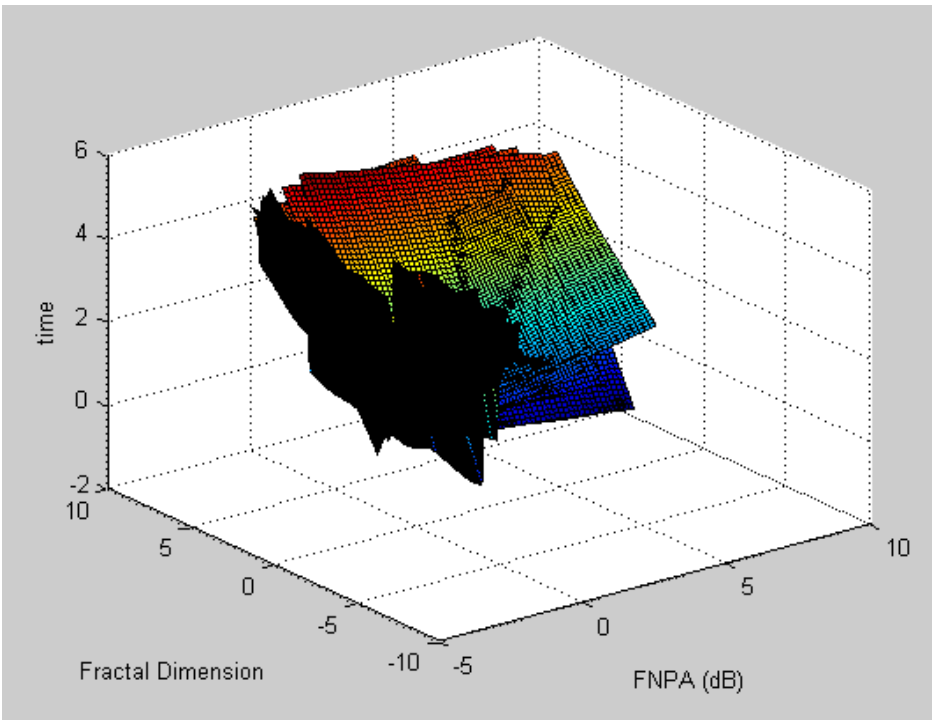


Fig. 5.6: Best-fit for all the 60 biopsy-proven patients, either malignant-proven or benign-proven, exhibiting 3 types of simulated over-time characteristic. These best-fits are linearly described by corresponding coefficients which are used for matching against new patients' feature changes during their follow-up checkups.

5.5 Relating different trends of physiological changes

From figure 5.5 it is obvious the amount of complexity that will add with thousands of more patients' data in trying to categorize the different trends of feature changes. This is similar to an expert's doubts in concluding upon one random patient's follow-up checkup based on his past experience of diagnosis of thousands of patients. Thus the need to sort trends of physiological changes becomes prominent. This work uses the fact that the best-fits illustrated in figure 5.6 are linearly described by corresponding displacement coefficients which can be used for matching against new patients' feature changes during their follow-up checkups. We avail this by computing correlation of a new patient's thin-plate spline (TPS) best-fit coefficients with all the past biopsy-proven patients' thin-plate spline (TPS) best-fit coefficients to quantify the trend of the new patient to different categories of previous biopsy-proven patients.

Table 5.2: Summary of different physiological changes simulated, taking in consideration of malignant tissues to exhibit poorer aspect ratio and margin definition, and higher value of FNPA and Fractal Dimension

Type	Characteristic Change
Benign Type A	Aspect Ratio ↑, Margin Definition ↑
Benign Type B (Non-Linear change)	Aspect Ratio almost same, Margin Definition ↓but later ↑ with time
Malignant Type C	Aspect Ratio ↓, Margin Definition ↓
Benign Type D	FNPA ↓, Fractal Dimension ↓
Benign Type E (Non-Linear Change)	FNPA almost same, Fractal Dimension ↑ but later ↓with time
Malignant Type F	FNPA ↑, Fractal Dimension ↑

5.6 Results

The three new patients' were simulated to have 5 sets of feature values, assuming a decision is made after fifth follow-up checkup. Quantification of matching of these data to simulated biopsy-proven data of sixty patients having 10 sets of feature values, assuming all were subject to 10 follow-up checkups, were evaluated in terms of strength of correlation. The following table shows strengths of correlation of the three new patients' over-time feature values against the 60 biopsy-proven patients.

Table 5.3: Result with changes related to Aspect ratio and Margin Definition.

		Previous patient's Data		
		Benign Type A	Benign Type B	Malignant Type C
New Patient's Data (Simulated)	Simulated Type A	**0.9493**	0.2637	-0.9202
	Simulated Type B	0.3554	**0.9123**	-0.3215
	Simulated Type C	-0.8863	-.4229	**0.8718**

Table 5.4: Result with changes related to FNPA (dB) and Fractal Dimension.

		Previous patient's Data		
		Benign Type D	Benign Type E	Malignant Type F
New Patient's Data (Simulated)	Simulated Type D	**0.9060**	0.4302	-0.8781
	Simulated Type E	0.1392	**0.8669**	-0.3594
	Simulated Type F	-0.8451	-0.3019	**0.8094**

5.7 Interpretation and robustness of the result

From the results observed, it is safe to say that this algorithm identifies patterns effectively. This is effectively similar to an expert looking into follow-up checkup images of a specific patient, and based on her knowledge about different patterns of changes she has seen, she makes a qualitative categorization of what type of change is being exhibited. Only in this case, the categorization is based on more tangible deductions and yielding a comparative correlation coefficient value to make better conclusions.

The phantom data were subjected to uniform and Gaussian distribution of changes. The changes were subject to variation of 100% to 300% of change, as depicted in the following single line instruction used in our simulation.

```
x(npat/3+1:npat*2/3,1)=0.4+.8*rand(npat/3,1)
```

Thus, it is safe to say that this algorithm is quite robust to non-linearity such as non-linear patient physiological changes, non-linear image readings from varying instruments, etc.

CHAPTER 6
DISCUSSIONS

6.1 Outlining Scope and Future Work

We have conclusively shown that our algorithm is far effective than normal hunch based decisions. This automation opens door to an effective community-wide program, in terms of its automation, more cost-effectiveness, lesser resource intensiveness and increased robustness. However, problems arise when real patient data is used instead of simulation, since the number of follow-up scans and time of scan can be widely varying. Thus, it may not be possible to categorize all types of over-time changes. Future works include collection of real-patients data and upgrading algorithm to yield conclusive findings of benign characteristics, with ideal zero percent of falsely concluding a malignant tissue to be benign.

6.2 Outlining a viable community-wide flow-chart for breast cancer diagnosis and treatment

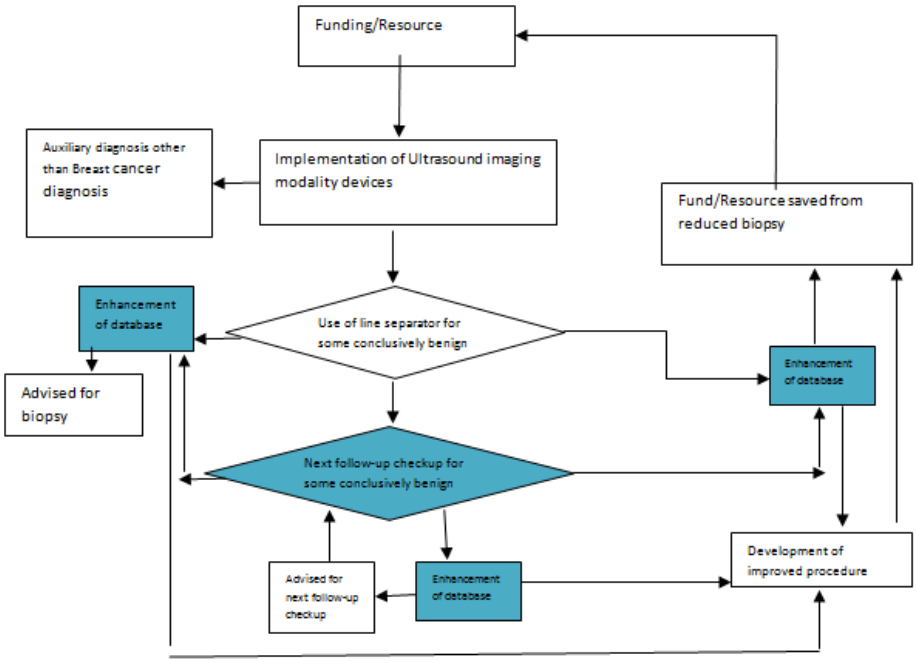


Fig. 6.1: Flow-chart outlining community-wide aspect.

The flow-chart in figure 6.1 illustrates how this automated follow-up checkup categorization comes in an overall scheme that can be implemented in fighting against breast cancer. This scheme is proposed to achieve a sustainable and well organized coordination across the whole community, which is the most crucial step in reducing mortality rates [1]. The use of computing devices and medical-telemetry added to our work will provide a well coordinated baseline that can eliminate a significant number of biopsies with minimal human intervention. This will make the whole process cheap and thereby sustainable. To add to it, improvement of database will naturally yield a better hypothesis in concluding a tissue characterization to be benign or malignant.

6.3 Problems faced

- The local diagnostic labs do not have organized ultrasound imaging data. Records of previously diagnosed data does not include later condition of that patient, i.e., whether she had conclusively benign or malignant lesion or other status.

- Although we have organized data for one particular exposures, we still lack data collected over time for specific patients. Thus we had to use an algorithm which is not yet limited to specific patterns.

6.4 Conclusion

A novel systematic quantization of follow-up checkups was demonstrated. We started with basic introduction of breast cancer and how ultrasound imaging modality is important in its diagnosis. BI-RADS criteria were then outlined and a quantization of acoustic and morphometric features of lesions was developed. Inconclusive quantization leads to follow-up checkups or biopsy depending on imminent risk factor of lesions. For follow-up checkups, rather than relying on educated guess of medical officers, we developed a systematic modeling and quantization technique using thin plate smoothing spline.

We have shown our technique to work effectively on four different sets of perturbations, each simulating specific sets of phantom physiological changes. Finally we have integrated our whole work into a scheme that aims at sustainable community-wide implementation for diagnosis of breast cancer. We showed how this integration will serve to reduce biopsy-costs incurred with simultaneous improvement of database and robustness of the algorithm.

6.5 Future Perspective

A copy of the results of this thesis is planned to be sent to Bangladesh Government, WHO and local NGO's. Once implemented, further studies on physiological changes and disease behavior may provide us conclusive insight. On a near-future note, we are working on getting data from a study on neoadjuvant therapy (or chemo-before-surgery). Results from this study should help us identify the physiological changes of lesions characteristic of neoadjuvant therapy.

REFERENCES

[1] WHO (2012)Breast cancer: prevention and control.

[2] WHO (2008). The global burden of disease: 2004 update.

[3] Coleman MP et al. (2008). Cancer survival in five continents: a worldwide population-based study (CONCORD). Lancet Oncol, 9, 730–56.

[4] Justo L. Sibala, MD; C. H. Joseph Chang, MD; Fritz Lin, MD; William R. Jewell, MD, Computed Tomographic MammographyDiagnosis of Mammographically and Clinically Occult Carcinoma of the Breast, Archives of Surgery . 1981;116(1):114-117.

[5] Yip CH et al. (2008). Guideline implementation for breast healthcare in low- and middle-income countries: early detection resource allocation. Cancer, 113, 2244–56.

[6] Vicini FA, Kestin L, Chen P, et al. Limited-field radiation therapy in the management of early-stage breast cancer, J National Cancer Insti tute 95, 1205-1210 (2003).

[7] Halliwell, M. 2010 A tutorial on ultrasonic physics and imaging techniques. J.Eng. Med.: Proc. IMechE 224, 127–142. (doi:10.1243/09544119JEIM656).

[8] Hasan Md. Kamrul, Anas Emran Mohammad Abu, Alam S. K., Lee Soo Yeol, Direct Mean Strain Estimation for Elastography Using Nearest-Neighbor Weighted Least-Squares Approach in the Frequency Domain, Ultrasound in Medicine and Biology, submitted, 2011.

[9] Frederic L. Lizzi, Michael Astor, Tian Liu, Cheri Deng, D. Jackson Coleman, Ronald H. Silverman, Ultrasonic Spectrum Analysis for Tissue Assays and Therapy Evaluation, International Journal of Imaging Systems and Technology, Special Issue: Acoustical Tomography, Volume 8, Issue 1, pages 3–10, 1997.

[10] Alam S.K., Feleppa EJ, Rondeau M, Kalisz A, Garra BS, Ultrasonic Multi-Feature Analysis Procedure for Computer-Aided Diagnosis of Solid Breast Lesions, Ultrason Imaging. 2011 Jan;33(1):17-38.

[11] American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) Atlas (American College of Radiology, Reston, VA, 2003).

[12] Mandelson EB, Berg NA, Merritt CRB. Towards a standardized breast ultrasound lexicon bi-rads: ultrasound, Semin Roentgenol 36, 217-25 (2001).

[13] Joo S, Yang YS, Moon WK, Kim HC. Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple monographic features, IEEE Trans Med Imaging 23, 1292–1300 (2004).

[14] http://www.scienceclarified.com/Ti-Vi/Tumor.html (2012).

[15] American College of Radiology (ACR) Breast Imaging Reporting and Data System Atlas (BI-RADS Atlas). Reston, Va: © American College of Radiology; 2003.

[16] Yao W, Zhao B, Zhao Y, Wang W, Qian G. Ultrasonographic texture analy sis of parenchymatous organs by the four-neighborhood-pixels algorithm, J Ultrasound Med 20, 465–471 (2001).

[17] Chou Y-H, Tiu C-M, Hung G-S, et al. Stepwise logistic regression analysis of tumor contour features for breast ultra sound diagnosis, Ultrasound Med Biol 27, 1493–1498 (2001).

[18] Hurst HE, Black RP, Simaika YM. Long Term Storage: An Experimental Study (Constable, London, 1965).

[19] Feder J. Fractals (Plenum Press, New York, 1988).

[20] Shankar PM, Dumane VA, Reid JM, et al. Classification of ultrasonic b-mode images of breast masses using nakagami distribution, IEEE Trans Ultrason Ferroelec Freq Contr 48, 569–580 (2001).

[21] Shankar PM, Dumane VA, Piccoli CW, et al. Computer-aided classification of breast masses in ultrasonic B-scans using a multiparameter approach, IEEE Trans Ultrason Ferroelec Freq Contr 50, 1002–1009 (2003).

[22] Drukker K, Giger ML, Mendelson EB. Computerized analysis of shadowing on breast ultrasound for improved lesion detection. Med Phys 30, 1833–42 (2003).

[23] Joo S, Yang YS, Moon WK, Kim HC. Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features, IEEE Trans Med Imag 23, 1292–1300 (2004).

[24] Wahba G, Spline models for observational data, Philadelphia: Society for Industrial and Applied Mathematics, 1990.

[25] Crochiere RE, Rabiner LR (1983). Multirate Digital Signal Processing. Englewood Cliffs, NJ: Prentice–Hall.

[26] http://www.mathworks.com/help/techdoc (2012).

APPENDIX A
SIMULATION SOFTWARE AND PROGRAMMING LANGUAGE

A.1 Sample coding provided for phantom data generation

%Generating scatter points

```
clear all
close all

npat=60;                %Number of old patients
nxy=11;                 %Number of samples from each patient
npat_new = 3;           %Number of new patients

x(1:npat/3,1)=0.4+.8*rand(npat/3,1);        %Starting points for benign type A
y(1:npat/3,1)=.05+.1*rand(npat/3,1);
z=zeros(npat,1);
figure (1)
scatter(x,y,'+')
hold on

x(npat/3+1:npat*2/3,1)=0.4+.8*rand(npat/3,1);          %Starting points for some
%sample malignant
y(npat/3+1:npat*2/3,1)=.05+.1*rand(npat/3,1);
scatter (x(npat/3+1:npat*2/3,1),y(npat/3+1:npat*2/3,1),'x')
hold on

x(npat*2/3+1:npat,1)=0.4+.8*rand(npat/3,1);         %Starting points for benign type
B
y(npat*2/3+1:npat,1)=.05+.1*rand(npat/3,1);
scatter (x(npat*2/3+1:npat,1),y(npat*2/3+1:npat,1),'*')
hold on
```

```
xlabel('Aspect ratio')
ylabel('Margin definition')

x_new(1:npat_new,1)=0.4+.8*rand(npat_new,1);          %Starting points for new patient
y_new(1:npat_new,1)=.05+.1*rand(npat_new,1);
z_new=zeros(npat_new,1);
scatter (x_new,y_new,'o')
```

%Next nxy-1 samples

```
for i=1:nxy-1

   x(1:npat/3, i+1) = x(1:npat/3, i)-0.2*rand(npat/3,1);          %Next nxy-1 points, nxy=number of samples per patient
   y(1:npat/3, i+1) = y(1:npat/3,i)+.25*rand(npat/3,1);           %AR dec, marg def inc [benign A]
   z(1:npat/3, i+1) = z(1:npat/3,i)+0.4+0.05*randn(npat/3,1);

   x_new(:, i+1) = x_new(1, i)-0.2*rand(3,1);
   y_new(:, i+1) = y_new(1,i)+.25*rand(3,1);
   z_new(:, i+1) = z_new(1,i)+0.4+0.05*randn(3,1);


   x(npat/3+1:npat*2/3, i+1) = x(npat/3+1:npat*2/3, i)+.2*rand(npat/3,1);
   y(npat/3+1:npat*2/3, i+1) = y(npat/3+1:npat*2/3, i)-.25*rand(npat/3,1); %AR inc, marg def dec  [malignant samples]
   z(npat/3+1:npat*2/3, i+1) = z(npat/3+1:npat*2/3, i)+0.4+0.05*randn(npat/3,1);

%     x_new(2, i+1) = x_new(2, i)+0.2*rand(1,1);
%     y_new(2, i+1) = y_new(2,i)-.25*rand(1,1);
```

```
%     z_new(2, i+1) = z_new(2,i)+0.4+0.05*randn(1,1);

   if i<nxy/2
      x(npat*2/3+1:npat, i+1) = x(npat*2/3+1:npat, i)+.02*rand(npat/3,1);
      y(npat*2/3+1:npat,     i+1)     =     y(npat*2/3+1:npat,     i)+.25*rand(npat/3,1);
%AR almost same, marg def inc     [benign B]
      z(npat*2/3+1:npat, i+1) = z(npat*2/3+1:npat, i)+0.4+0.05*randn(npat/3,1);



%       x_new(:, i+1) = x_new(:, i)+.02*rand(3,1);
%       y_new(:, i+1) = y_new(:, i)+.25*rand(3,1);
%       z_new(:, i+1) = z_new(:, i)+0.4+0.05*randn(3,1);

   else

      x(npat*2/3+1:npat, i+1) = x(npat*2/3+1:npat, i)-.2*rand(npat/3,1);
      y(npat*2/3+1:npat,     i+1)     =     y(npat*2/3+1:npat,     i)+.25*rand(npat/3,1);
%AR dec, marg def inc     [benign B continued]
      z(npat*2/3+1:npat, i+1) = z(npat*2/3+1:npat, i)+0.4+0.05*randn(npat/3,1);

%       x_new(:, i+1) = x_new(:, i)-.02*rand(3,1);
%       y_new(:, i+1) = y_new(:, i)+.25*rand(3,1);                %New patient x1, y1,
z1 Remember !!!
%       z_new(:, i+1) = z_new(:, i)+0.4+0.05*randn(3,1);
   end

end
```

A.2 Simulation software used: MATLAB 2010 TM

MathWorks is the world's leading developer of technical computing software for engineers and scientists in industry, government, and education.

MATLAB (matrix laboratory) is a numerical computing environment and fourth-generation programming language. Developed by MathWorks, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, and Fortran.

Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPAD symbolic engine, allowing access to symbolic computing capabilities. An additional package, Simulink, adds graphical multi-domain simulation and Model-Based Design for dynamic and embedded systems.

In regard of this thesis, the importance of MATLAB comes in generating and processing stochastic data. Most crucial of all, is the availability of powerful curve-fitting and 3-D curve generation techniques. An example of a 3-D curve-genearation is shown in figure A.1 [26].

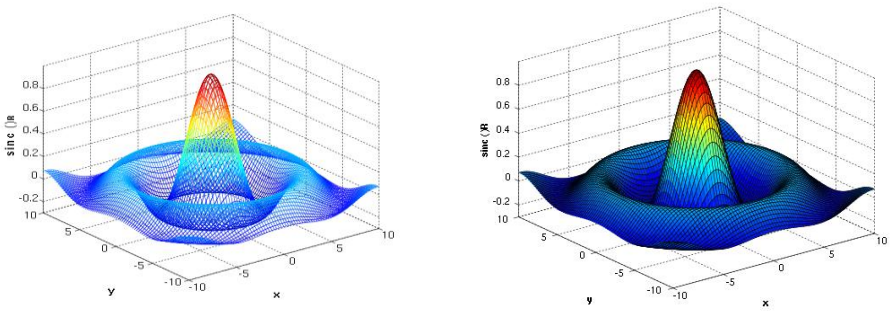


Fig. A.1: A wireframe 3D and surface 3D plot of the two-dimensional un-normalized sinc function