FEATURE EXTRACTION FROM MAMMOGRAPHIC MASS SHAPES AND DEVELOPMENT OF A MAMMOGRAM DATABASE

by

Gökhan Ertaş

B.S. in Electrical and Electronics Engineering,

Erciyes University, 1999

Submitted to the Institute of Biomedical Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Biomedical Engineering

Boğaziçi University

June, 2001

FEATURE EXTRACTION FROM MAMMOGRAPHIC MASS SHAPES AND DEVELOPMENT OF A MAMMOGRAM DATABASE

APPROVED BY:

Assoc. Prof. Dr. (Thesis Supervisor)	Halil Özcan Gülçür	.....
Assoc. Prof. Dr.	Albert Güveniş	.....
Assoc. Prof. Dr.	Erkin Arıbal	.....

DATE OF APPROVAL: June 11, 2001

ACKNOWLEDGMENTS

I am deeply grateful to my thesis supervisor, Assoc. Prof. Dr. H. Özcan GÜLÇÜR for his academic support and friendly tolerance. I also deeply appreciate the invaluable assistance provided by Assoc. Prof. Dr. Erkin ARIBAL. I am extremely grateful for his practice and collaboration.

Very special thanks are to be extended to Dr. Aslı SEMİZ, who was very friendly while teaching me mammography and sonography, and helping me in searching biopsy proven mammographic patient records.

FEATURE EXTRACTION FROM MAMMOGRAPHIC MASS SHAPES AND DEVELOPMENT OF A MAMMOGRAM DATABASE

ABSTRACT

Breast cancer is one of the most common malignancies in women and a rare malignancy in men. It has been widely reported that breast cancer has become the second leading cause of cancer death among women. Over a lifetime, one in nine women risk contracting breast cancer. However, women who are diagnosed at an early stage can survive this often deadly disease. Mammography provides the best screening modality for detecting early breast cancer, even before a lesion is palpable. Because of the malignant mass pathology, the shape of the mammographic mass can be used to discriminate between malignant and benign masses.

In this study geometric parameters such as area, perimeter, circularity, normalized circularity, radial distance mean and standard deviation, area ratio, orientation, eccentricity, moment invariants and Fourier descriptors up to ten, are calculated. The process starts with a segmentation phase, in which an expert radiologist segments the mammographic mass shapes within the mammographic database set. These pre-segmented mammographic mass shapes are then processed by a mass boundary detection algorithm to obtain descriptive geometric parameters. A carefully designed classification scheme is used in the final step to classify masses as benign or malign. The results show that normalized circulatory area and the Fourier descriptors can be used successfully for feature extraction. The software developed utilizes this finding in the automatic classification of the suspicious masses.

A mammogram database designed to store the images of the masses, calculated shape descriptor parameters and some additional data, such as patient history, category of the mass and biopsy report, if performed, which are required in BI-RADS is also introduced. The developed database is designed to be an Open Database Connectivity compliant relational database to support some future uses, such as screening the growth of suspicious masses, telemedical service support for sharing mass information and for facilitating statistical data analysis. A touch on memory system has been used as a tool to permit secure access to the electronic patient record in the mammogram database. The software is written in Delphi and runs on machines equipped with MS Windows.

Keywords: Breast Cancer, Mammography, Bayesian Classifier, ROC Analysis, Relational Database Management System, Touch on Memory, BI-RADS

MAMOGRAFİK KÜTLE ŞEKİLLERİNDEN ÖZELLİK ÇIKARIMI VE BİR MAMMOGRAM VERİ TABANI GELİŞTİRİLMESİ

ÖZET

Meme kanseri kadınlarda yaygın erkeklerde ise nadiren görülen kanserlerden biridir. Kadınlar arasında meme kanserinin ikinci kanser sebepli ölüm sebebi olduğu rapor edilmektedir. Bir yaşam süresi boyunca her dokuz kadından biri meme kanseri riski ile karşılaşmaktadır. Bununla birlikte erken aşamada teşhis edilen kadınlar bu ölümcül hastalıktan sağ olarak kurtulmaktadır. Mammografi erken meme kanseri tespitinde, yani lezyon ele gelir hale gelmeden önce, en iyi görüntüleme modalitesidir. Kötü huylu kütle patolojisinden dolayı mammografik kütle şekilleri kötü huylu ve iyi huylu kütleleri ayırt etmek için kullanılabilmektedir.

Bu çalışmada geometrik parametreler olarak alan, perimetre, yuvarlaksallık, radyal mesafe ortalaması ve standart sapma, alan oranı, yönlendirme, ayrıksallık, moment değişmeyenleri ve Fourier betimleyicileri hesaplanmaktadır. İşlem bir kesimleme aşamasıyla başlar ki burada radyoloji uzmanı mammografi setindeki mammografik kütle şekillerini kesimler. Bu ön kesimlenmiş mammografik kütle şekilleri daha sonra, betimsel geometrik parametrelerin elde edilmesi için, bir kütle çeperi tespit algoritmasıyla işlenir. Titizlikle tasarlanmış bir sınıflandırma planı son aşamada kütlelerin kötü huylu veya iyi huylu olarak sınıflandırılması için kullanılır. Sonuçlar göstermektedir ki normalize yuvarlaksallık, alan ve Fourier bitimleyicileri özellik çıkarımında başarıyla kullanılabilmektedirler. Geliştirilen yazılım bu bulgudan şüpheli kütlelerin otomatik sınıflandırılmasında yararlanmaktadır.

Bir mammogram veri tabanı kütle görüntülerini, hesaplanmış şekil betimleyici parametreleri, BI-RADS' ta ihtiyaç duyulan hasta geçmiş, kütle kategorisi ve yapılmışsa biyopsi raporu gibi ek bilgileri saklamak için tasarlanmıştır. Geliştirilen veri tabanı şüpheli kütlelerin büyümesin gözlenmesi, kütle bilgi paylaşımı ve istatistiksel veri analizi için tele tıp desteği gibi gelecek kullanımlara destek vermek için açık veri tabanı bağlanabilirliği uyumlu bir ilişkisel veri tabanı olacak şekilde tasarlanmıştır. Bir Touch on Memory sistemi mammogram veri tabanındaki elektronik hasta verilerine güvenli erişime imkan tanıyan bir araç olarak kullanılmıştır. Yazılım Delphi de yazılmıştır ve MS Windows yüklü makinelerde çalışmaktadır.

Anahtar Kelimeler: Meme Kanseri, Mammografi, Bayesian Sınıflandırıcı, ROC Analizi, İlişkisel Veri Tabanı Yönetim Sistemi, Touch on Memory, BI-RADS

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii
ABSTRACT iv
ÖZET v
TABLE OF CONTENTS vi
LIST OF FIGURES x
LIST OF TABLES xii
LIST OF SYMBOLS xiii
1. INTRODUCTION 1
1.1 Motivation, Background and Objectives 1
1.2 Block Diagram of the Developed System 2
1.3 Outline of the Thesis 2
2. BREAST CANCER 4
2.1 Introduction 4
2.2 Anatomy of Breast 8
2.3 Risk Factors and Medical Patient History 10
2.4 Physical Examination of Breast 14
2.5 Breast Imaging Modalities 15
2.5.1 Mammography 15
2.5.2 Sonography 15
2.5.3 Contrast Enhanced MRI 19
3. MAMMOGRAPHY 23
3.1 Introduction 23
3.2 Components of Mammographic Imaging 23
3.2.1 X-Ray Tube 23
3.2.2 Sharpness 25
3.2.3 Radiation Spectrum 25
3.2.4 Target / Filter Combination 26
3.2.5 Peak Kilovoltage (kVp) 27
3.2.6 Penetration: The Heel Effect 28
3.2.7 Scattered Radiation 28

3.2.8 Scattered Reduction 28
3.3 Accuracy 29
3.3.1 Sensitivity 29
3.3.2 Specificity 30
3.4 Compression and Positioning 31
3.4.1 Compression 31
3.4.2 Positioning for Standard Views 33
3.4.3 Mediolateral Oblique View 33
3.4.4 Craniocaudal View 35
3.5 ACR Standard for the Performance of Diagnostic Mammography 35
3.6 ACR Standard For The Performance Of Screening Mammography 39
4. MAMMOGRAPHIC MASS SHAPE ANALYSIS 41
4.1 Introduction 41
4.2 Simple Geometric Parameters 44
4.2.1 Area 44
4.2.2 Centroid 44
4.2.3 Perimeter 45
4.2.4 Normalized Circularity 45
4.2.5 Radial Distance, Mean and Standard Deviation 46
4.2.6 Area Ratio 46
4.3 Moment Based Shape Features 47
4.3.1 Orientation 48
4.3.2 Eccentricity 48
4.3.3 Moment Invariants 49
4.4 Fourier Descriptors 50
5. CLASSIFYING MAMMOGRAPHIC MASS SHAPES 52
5.1 Introduction 52
5.2 Feature Extraction 53
5.3 Classification Scheme 53
5.3.1 Bayes Theorem and Bayes Risk Criterion 58
5.3.2 ROC Analysis 61
5.3.3 Classification Scheme Performance 63

6. MAMMOGRAM DATABASE DEVELOPMENT 65
6.1 Introduction 65
6.2 Relational Database Management System 65
6.3 Structured Query Language (SQL) 67
6.4 Open Database Connectivity (ODBC) 69
6.5 Designing Mammogram Database 70
6.6 Patient File Table 70
6.7 Reports Table 72
6.8 Findings Table 73
6.9 Biopsies Table 74
6.10 Doctors Table 74
6.11 Assessments Table 74
6.12 Recommendations Table 75
6.13 Reasons Table 76
6.14 BiopsyTech Table 77
6.15 Forms 78
6.16 Patient Privacy 84
7. AUTOMATIC PATIENT IDENTIFICATION 85
7.1 Introduction 85
7.2 DS1990A Serial Number iButton 86
7.3 Hardware Configuration 87
7.4 DS9097A Com Port Adapter 96
8. CONCLUSIONS 99
8.1 Introduction 99
8.2 Results 99
8.3 Recommendations for Future Work 100
APPENDIX A 101
THE ACR BREAST IMAGING REPORTING AND DATA SYSTEM (BI-RADS) 101
A.1 Introduction 101
A.2 Breast Imaging Lexion 101
A.3 Reporting System 108
A.4 Follow-Up and Outcome Monitoring 112

A.5 ACR National Mammography Database 112
REFERENCES 113

LIST OF FIGURES

Figure 1.1 Block diagram of the developed system.....2
Figure 2.1 Breast cancer spread in its early and late stage.....6
Figure 2.2 Schematic representation of the basic anatomy of the breast.....9
Figure 2.3 Orientation of pectoralis major and minor muscles.....9
Figure 2.4 Cancers are thought to arise in the intralobular portion of the terminal duct....11
Figure 2.5.a Incidence curve depending on patient age.....11
Figure 2.5.b Logarithmic plot of incidence curve.....12
Figure 2.6 The age of menarche. This seems to get earlier and earlier.....13
Figure 2.7.a Normal case, patient age 66, digitizer dba 21.....16
Figure 2.7.b Benign case, patient age 39, digitizer lumisys laser.....16
Figure 2.7.c Malign or cancer case, patient age 85, digitizer dba 21.....17
Figure 2.8 Mammogram and ultrasound view of a breast mass.....18
Figure 2.9 Image of a benign breast tumor.....18
Figure 2.10 Image of a ductal invasion associated with malignant breast mass.....18
Figure 2.11 A 2-D section of an MRI image of a female breast.....19
Figure 2.12.a MR-Mammogram before contrast injection.....21
Figure 2.12.b MR-Mammogram after contrast injection.....21
Figure 2.12.c Subtracted image, only the cancer is visible.....22
Figure 3.1. The components of X-ray device.....24
Figure 3.2 Conventional mammography.....24
Figure 3.3 Digital mammography.....25
Figure 3.4.a A well-positioned mediolateral oblique projection.....34
Figure 3.4.b Mammogram.....34
Figure 3.5 A craniocaudal projection.....35
Figure 4.1 Morphologic spectrum of mammographic masses.....41
Figure 4.2.a A nodular benign tumor.....42
Figure 4.2.b A lobular benign tumor.....42
Figure 4.2.c A lobular malign tumor.....43
Figure 4.2.d A stellate malign tumor.....43
Figure 4.3 A digital boundary and its representation as a complex sequence.....50
Figure 4.4 Normalization for Fourier descriptors.....51

Figure 5.1.a Statistical pattern recognition approach. 55
Figure 5.1.b Syntactic pattern recognition approach. 55
Figure 5.1.c Neural network approach 55
Figure 5.2 Densities for a single measurement. 57
Figure 5.3 Converting $P(w_i)$ to $P(w_i \mid x)$ via x. 58
Figure 5.4 The decision tree for classification. 61
Figure 5.5 Different ROC curves. 62
Figure 6.1 Codd' s relational database management system 66
Figure 6.2 ODBC structure. 69
Figure 6.3 Relationships of tables in mammogram database. 71
Figure 6.4 Database login form. 79
Figure 6.5 The case in which patient record could not found 79
Figure 6.6 New patient data entry form. 80
Figure 6.7 Patient file form. 82
Figure 6.8 Patient file form after a new mass entry. 83
Figure 7.1 DS1990A memory map. 86
Figure 7.2 Physical dimensions of DS1990A F5 87
Figure 7.3 DS1990A equivalent circuit. 88
Figure 7.4 Standard TTL as bus master. 88
Figure 7.5 Initialization procedure, reset and presence pulses 90
Figure 7.6.a Read / write timing diagram, write-one time slot. 91
Figure 7.6.b Read / Write timing diagram, write-zero time slot 92
Figure 7.6.c Read / Write timing diagram, read .data time slot. 92
Figure 7.7 CRC generated using a polynomial generator consisting of a shift register and XOR gates 93
Figure 7.8 Interface between TTL and 1-wire system. 94
Figure 7.9 8250 UART block diagram. 94
Figure 7.10 DS9097A com port reader as a iButton reader 97
Figure 7.11 Schematic layout of DS9097A. 98

LIST OF TABLES

Table 4.1 Basic Properties of Fourier Descriptors 51

Table 5.1 Traditional Academic Point System for Accuracy 63

Table 5.2 Decisions of The Software 64

Table 6.1 Codes and Descriptions for Assessments as Described in BI-RADS 75

Table 6.2 Codes for Recommendations as Defined in BI-RADS 76

Table 6.3 Reason Codes and Description for Each Code as Defined in BI-RADS 77

Table 6.4 Biopsy Technique Codes and Descriptions Defined in BI-RADS 78

Table 7.1 Read and Search ROM Commands 89

Table 7.2 RS232 Connections for 9 and 25 Pins. 96

LIST OF SYMBOLS

A	Area of the mammographic mass shape
A_{ratio}	Area ratio
$a(u)$	Discrete Fourier Transform of s(i)
B	Bayesian cost
c	Circularity
C_{jk}	Cost of deciding d_j when m_k is true
c_N	Normalized circularity
$d(i)$	Radial distance
d_{avg}	Radial distance mean
eccentricity	Eccentricity of a region
$M_{p,q}$	Moments of order p+q
p	Perimeter
$P(w_i)$	A priori class probability
$P(w_i \mid \underline{x})$	A posteriori probability
$s(i)$	Complex representation of a point in x, y axis
(x_c, y_c)	Centroid of the mass
σ	Standard deviation of radial distance
$\mu_{p,q}$	Central moments of order p+q
λ	Principal moments
θ	Orientation of a region
$\eta_{p,q}$	Normalized central moments

1. INTRODUCTION

1.1 Motivation, Background and Objectives

Breast cancer is one of the most common malignancies in women and a rare malignancy in men. It has been widely reported that breast cancer has become the second leading cause of cancer death among women. Over a lifetime, one in nine women risk contracting breast cancer. However, the good news is that women who are diagnosed at an early stage can survive this often-deadly disease. The risk of breast cancer tends to increase with age.

Mammography, which is an X-ray examination of the breasts, provides the best screening modality we have at this time for detecting early breast cancer, even before a lesion is palpable. It uses safe, low doses of radiation in the form of X-rays to image the inside of the breast. The X-rays pass through the breast and create an image on film. A radiologist, a physician who specializes in medical diagnosis by X-ray, examines the mammogram to detect any variation from the norm. Approximately 80 to 85% of localized breast cancers are diagnosed by the mammographic appearances of the tumor [1].

In this thesis a computer system developed for breast cancer detection is introduced. The system is based on mammographic mass shape analysis. These shape features are useful for classification systems to classify masses as round, nodular, or stellate. After the feature extraction from segmented mammographic masses, shape descriptor parameters are stored in a well designed and Open Database Connectivity (ODBC) compliant relational database for future use and telemedical service support.

The security of the mammography database and the electronic record privacy of the patient are guarantied by employing a Touch Memory system and an encryption algorithm. The software developed includes modules for mammography database management, telemedicine applications and statistical data analysis..

1.2 Block Diagram of the Developed System

A block diagram of the developed system is shown in Figure 1.1. The computational results of shape descriptors show that normalized circulatory area and the Fourier descriptors can be used successfully for feature extraction. In addition to this, it is shown that the developed mammogram database meets the requirements to query processes for statistical data analysis while touch on memories are suitable for automatic patient identification and privacy. An outline of the thesis is given in the next section.

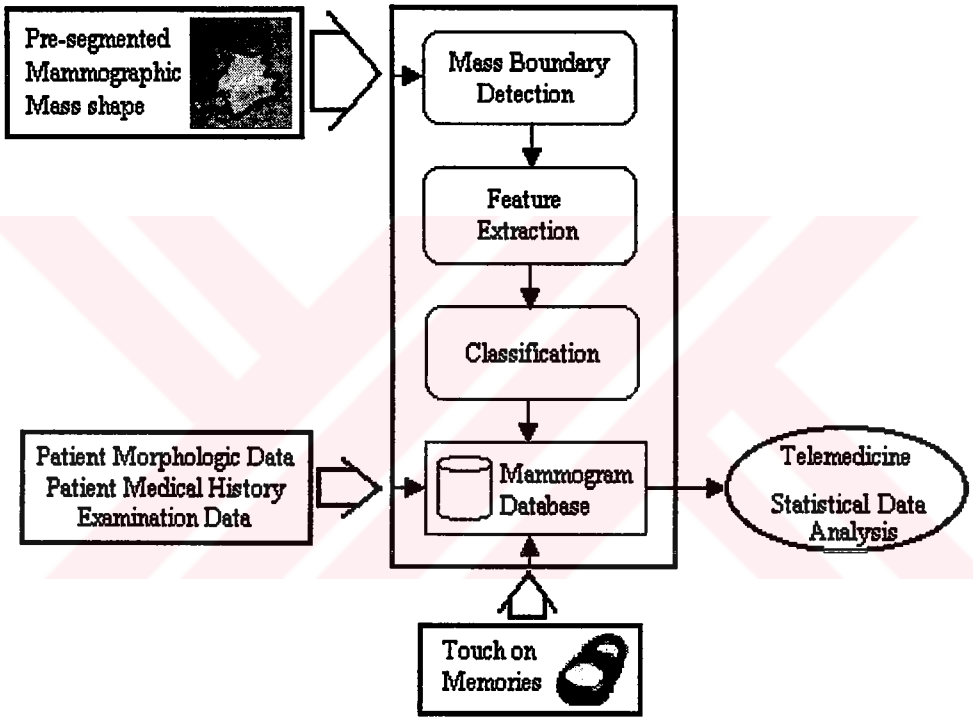

Figure 1.1 Block diagram of the developed system.

1.3 Outline of the Thesis

Chapter 2 covers the anatomy of breast, importance of breast cancer, causes of breast cancer, risk factors and patient history, physical examination and common breast imaging modalities used to detect malignancy. In Chapter 3, mammography, as a breast imaging modality, which provides the best screening modality for detecting early breast cancer, even before a lesion is palpable, is explained in details. The components of mammographic imaging technique, accuracy, and importance of positioning and

compression, diagnostic and screening mammography are the covered topics. Primary features of malignancy and the ability to discriminate between malignant and benign masses with respect to their shape properties are proposed in Chapter 4. The shape descriptors and computation methods used in the developed software are the other subjects of this chapter. Chapter 5 describes the application of shape analysis to computerized mammographic methods and explains a Bayesian classifier based classification scheme design. ROC analysis, that answers the reasons of selecting the shape descriptors used in classification, is also defined in this chapter. Breast imaging reporting and data system, BI-RADS, is focused in Chapter 6. This system is a quality assurance tool designed to standardize mammographic reporting, reduce confusion in breast imaging interpretations, and facilitate outcome monitoring. Through a medical audit and outcome monitoring, the system provides important peer review and quality assurance data to improve the quality of patient care. The requirements of relational databases, a well designed relational mammography database structure and views of the system software forms are discussed in Chapter 7. Touch on memories, as automatic identification tools and the idea of employing this technology to patient privacy for the mammogram database are the subjects introduced in Chapter 8.

2. BREAST CANCER

2.1 Introduction

Although lung cancer has surpassed breast cancer as the leading cause of cancer death among women, breast cancer remains the leading cause of non-preventable cancer death. On the basis of gender alone, all women have a baseline risk of developing breast cancer of approximately 4% to 6% over the course of their lifetime. A sub-segment of the population is at additional risk, giving the overall population an average risk of one chance in nine (11%) of developing breast cancer by the age of 85.

Breast cancer is one of the best-studied human tumors, but it remains poorly understood. It has become fairly certain that as with all solid tumors, breast cancer is the result of DNA alterations (damage or mutation) that lead to uncontrolled cell proliferation. Nevertheless its actual etiology remains obscure. It is not possible to predict who will develop breast cancer. Methods to prevent it are under investigation but remain unproved while methods to cure it are controversial and unfortunately not always successful Between 30% and 40% of women who develop breast cancer die from it.

Despite the numerous large studies that have evaluated the benefit of early detection, there is no universal agreement as to who should be screened, at what age screening should begin, how often an individual should be screened and at what age screening should be discontinued.

Medical understanding of breast anatomy, histology and pathology remains rudimentary and basic questions remain unanswered. These range from simplest anatomic questions such as the distribution of ducts and their interrelation, to more complex questions such as the possibility that there is a stem cell that is responsible for terminal duct differentiation and possibly the site for malignant transformation.

There is no general agreement as to whether breast cancer originates as a disease of a single cell whose progeny spread through a single duct system or as a field phenomenon in which multiple cells and ducts (related or separate) are simultaneously involved. It is not known why the cells of the terminal duct appear to be the site of origin for breast cancer and there is no basic understanding of the relationships between cancers that appear to originate from cells lining the duct and those that appear to originate from cells that line the lobule. There are clearly lesions such as a typical hyperplasia and lobular carcinoma in situ (lobular neoplasia) that increase the risk of developing cancer but the relationships with invasive cancer remains obscure. Even ductal carcinoma in situ which has direct links to invasive cancer, remains enigmatic since it appears that not all ductal carcinoma in situ progresses to invasive cancers.

The complexity and frustration of breast cancer are reflected in the fact that some women, with microscopic or even undetectable, breast cancer die rapidly of metastatic disease while other women with large masses that almost replace the breast and involve axillary lymph nodes survive many years. Fortunately, a disease process does not have to be understood to be treated successfully. The randomized controlled trails of screening have clearly demonstrated that the natural history of breast cancer can be interrupted if the process is detected early enough and cure or delayed mortality can be achieved for many women. The efficacy of screening asymptotic women has been demonstrated particularly through the use of mammography (Figure 2.1). Mammography is the primary breast imaging technology and the only system that has been validated for screening. Nevertheless other tests have been and will be shown to be valuable in the assessment of breast problems.

Since breast imaging was first described as a specialty within radiology, the field has expanded to include virtually all of the imaging technologies including ultrasound and magnetic resonance imaging [2]. The specialty attempts using these various methods of imaging breast to detect and diagnose breast cancer earlier when cure is more likely and treatment less formidable and to assist in determining the appropriate therapy for a given lesion.

Earlier detection has permitted the more widespread use of conservation therapy with the excision of primary lesion and radiation to reduce the likelihood of the recurrence

within breast. When found in its earliest stage, breast cancer is 90 percent curable. As techniques are refined, it may be possible to eliminate the tumor within breast with little physical damage to the normal breast tissues.

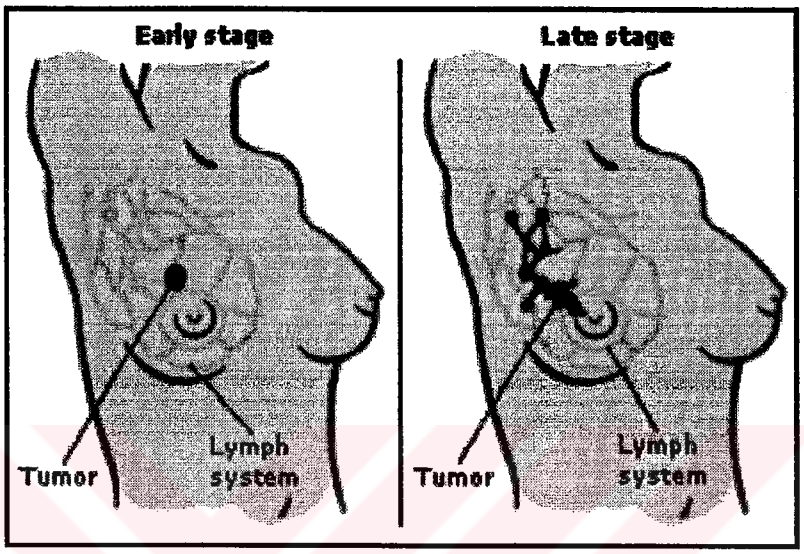

Figure 2.1 Breast cancer spread in its early and late stage.

Within the specialty, there are numerous gaps in medical knowledge. The very basic natural development of the breast as manifest by its mammographic appearance is poorly understood. Is there a normal breast? There has never been a study to determine whether or not there are significant changes on the mammogram that are related to the menstrual cycle. The histologic basis of all the structures seen on the mammographic images is still not clear. Why does one woman have dense breast tissue while another, of the same age and apparent demographic characteristics, have predominantly fat? Is one of these a normal state while the other is an indicator of abnormality? There has been a fairly consistent suggestion that women with more radio-graphically dense tissues are at a slightly increased risk for cancer but the relationships and mechanisms for this remain obscure.

Screening has been shown to be efficacious but there are strong disagreements as to the appropriate methods of delivery of mammographic services. Is physical examination necessary? Who should perform it? How involved must the radiologist be? Is

interpretation of the screening study sufficient or must the radiologist examine and correlate the clinical examination with the mammogram? Does the psychological benefit of giving the patients an immediate report outweigh the increased detection of cancers afforded by a delayed double reading? What is the role of the radiologist in the evaluation of the palpable abnormality? Additional projections and spot compression views reveal more cancers on mammogram? But does this actually alter management? Similarly there is disagreement over the role of ultrasound in the evaluation of the palpable mass. The indiscriminate use of ultrasound merely adds to the expense if needle aspiration is to be undertaken to simultaneously diagnose and eliminate a palpable cyst. What should the threshold for intervention be when an abnormality is seen on a mammogram? [3-6]. These and many other questions remain unsolved. Most should be answered through scientific study. For some the answer is philosophical while the economics of health care may ultimately dictate many choices [7].

The heightened awareness of the breast cancer problem has simulated a renewed effort and increased research support to try to understand, detect, treat and one day prevent breast cancer. The BRCA1 and BRCA2 genes, found in women with hereditary breast cancer, have been isolated. Women with an abnormal copy of BRCA1 or BRCA2 genes face a 50% lifetime risk of developing breast cancer. Although the inheritance of a faulty BRCA1 or BRCA2 gene accounts for only 10% of all cancers, an understanding of the role of these genes in the development of breast cancer may point the direction toward more effective approaches to detection, treatment or prevention.

There is great optimism over the potential of molecular biological research and unraveling the molecular basis of cancers will likely lead to more effective approaches to the problem, but there have been high hopes for basic research in the past, only to end in disappointment. Clinical research is of equal importance and should be strongly supported. There is no question that mammographic screening is not the solution to the problem of breast cancer, but it is the best that is available at the present and for the foreseeable future. If properly performed, mammographic screening can reduce the death rate by at least 25% to 30% and probably more. Attention should be given to its proper performance [1].

2.2 Anatomy of Breast

A basic understanding of the anatomy and histology of the breast and of the complex underlying microscopic structures, in which changes take place, is important for an understanding of the pathologic processes that occur and is helpful for image interpretation. The breast is a modified skin gland. It develops on the chest wall between the clavicle and the sixth to eight ribs. Breast tissue can be found as far medially as the sternum and laterally to the midaxillary line. Breast tissue frequently extends around the lateral margin of the pectoralis major muscle and may be found high in the axilla, occasionally reaching to its apex.

The skin of the breast is usually 0.5 to 2 mm in thickness. Just beneath the skin lies the superficial layer of fascia that, at the level of the breast, divides into superficial and deep layers as seen in Figure 2.2. The breast develops between this split layer of fascia and is enveloped by it. The deep layer of this split fascia forms the retro-mammary fascia and this lies immediately on the fascia that overlies the pectoralis major muscle providing surfaces that permit some movement of the breast on the chess wall. The fascial layers do not completely isolate the breast from the pectoralis major. Blood vessels and lymphatics penetrate the fascial layers, coursing between the muscle and the breast.

The structure of the pectoralis major muscle is important for breast imaging as seen in Figure 2.3. The mediolateral oblique projection is positioned so that the plane of compression is parallel to the oblique fibers of the free margin of the muscle as it extends from the ribs to the humerus. This permits maximum traction on the breast so that it can be fully positioned over the detector and comfortably compressed.

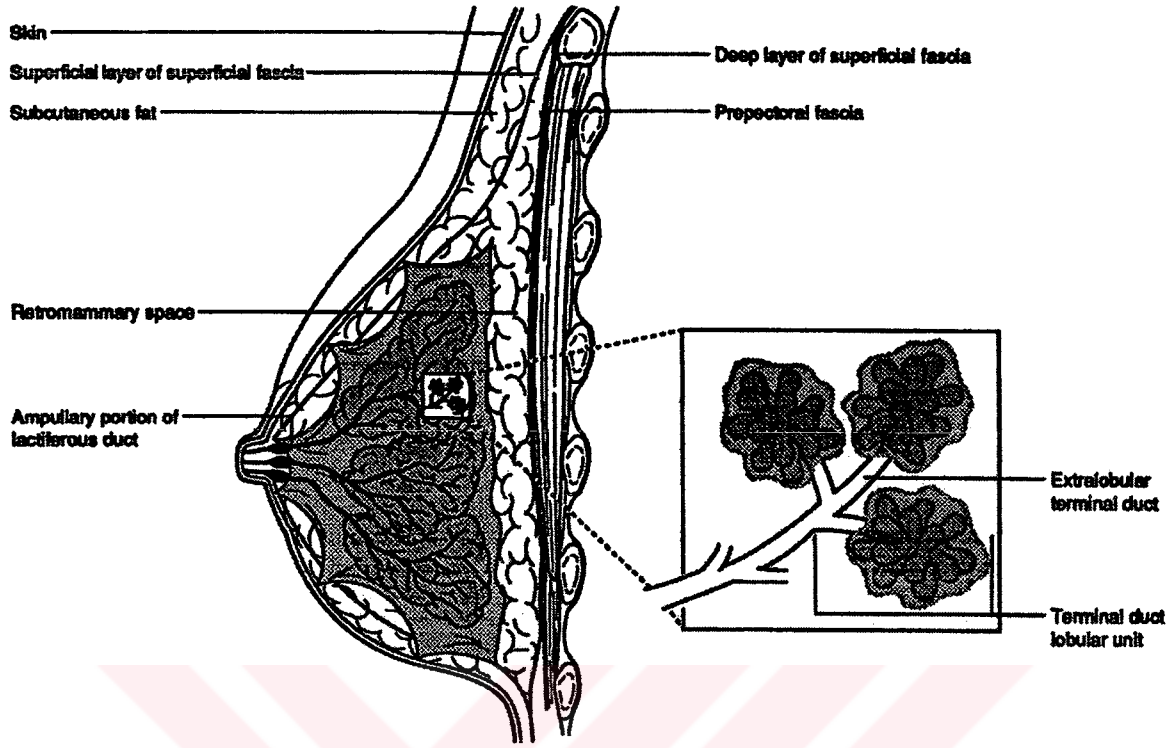

Figure 2.2 Schematic representation of the basic anatomy of the breast.

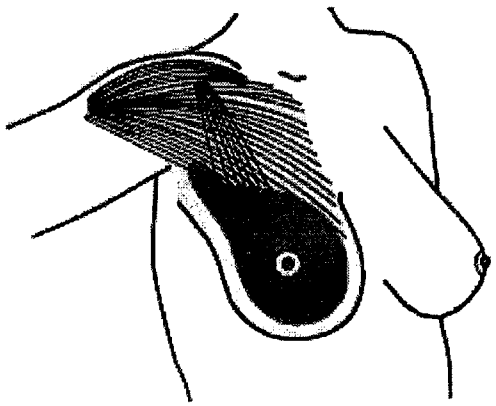

Figure 2.3 Orientation of pectoralis major and minor muscles.

The breast is divided into incomplete compartments by varying amounts of connective tissue, described by Cooper 150 years ago and known as Cooper' s ligaments. These planes of collagen come to peaks that attach to the skin as the retinacula cutis. The breast is primarily supported by the skin and is anchored to the chest wall medially along the sternum and superiorly toward the clavicle. It is fairly mobile inferiorly and along its lateral margins. The skin of the breast resists displacement downward but permits elevation. Displacement laterally is also resisted but the breast is fairly easily moved medially. The movements are used to advantage in mammographic positioning.

Immediately beneath the skin is subcutaneous fat. This layer varies with individuals. In some women it is clearly separate from the parenchymal cone of the breast. In others the subcutaneous fat cannot be distinguished from the fat between the glandular structures.

The nipple contains many sensory nerve endings and smooth muscle bundles. The surface of the nipple itself irregular and contains numerous crevices. The duct orifices are at the bottom of these crevices. Various investigators suggest that there are between 8 and 20 major ducts open on the nipple. Each of these ducts and its tributaries defines a lobe or segment of the gland. Beneath the nipple openings, the major ducts dilate into their ampullary portions. These are lactiferous sinuses. The deeper segmental ducts divide into sub-segmental structures and may branch further until they form the terminal duct that enters the lobule. Terminal ducts are the places where cancers are thought to develop as seen in Figure 2.4.

2.3 Risk Factors and Medical Patient History

The most important risk factors are age, family history, personnel history of the patient, immunodeficiency disorders, early menarche or late menopause, breast-feeding frequency and duration. The incidence of breast cancer begins to increase around age 35, has a slight plateau around age 50 and then continues to increase again at a more gradual slope suggesting to different populations of cancer as seen in Figure 2.5.a. The change in slope can be seen when the data are graphed on a logarithmic plot. There is a rapid rise in incidence until around the age of 50 and then the rate of increase in incidence slows as seen in Figure 2.5.b.

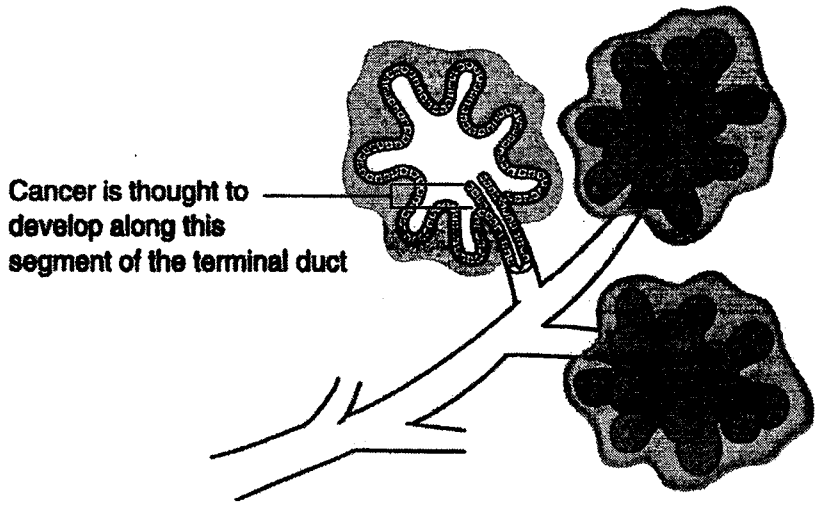

Figure 2.4 Cancers are thought to arise in the intralobular portion of the terminal duct.

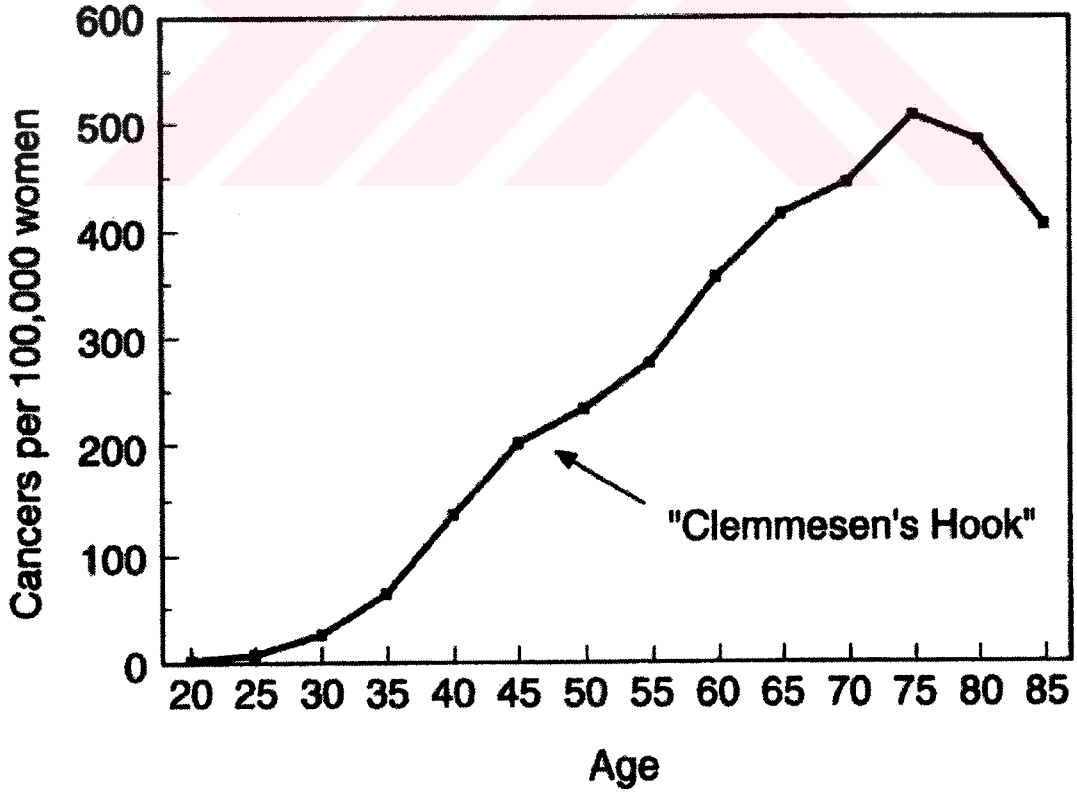

Figure 2.5.a Incidence curve depending on patient age

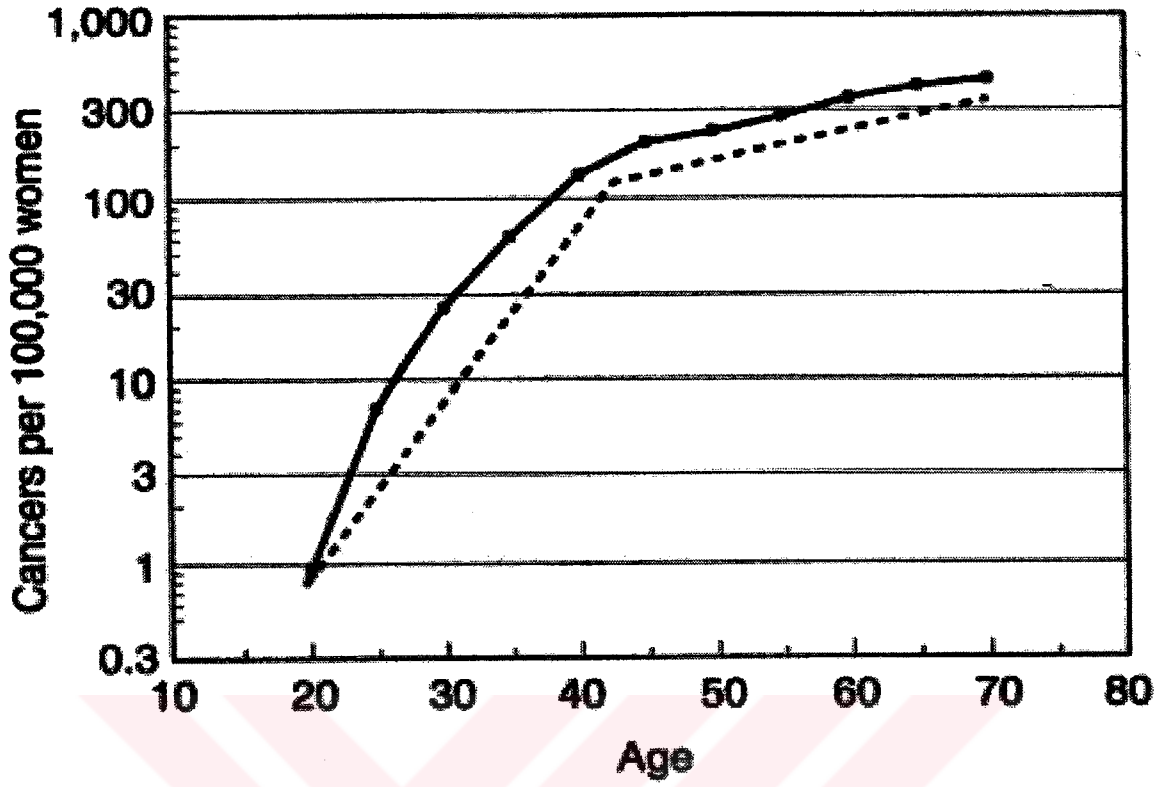

Figure 2.5.b Logarithmic plot of incidence curve.

A history of breast cancer in immediate relatives, the number of family members affected and the time at which the cancer appeared (premenopausal or postmenopausal) are significant [8]. The patient's personnel history of an invasive or in situ breast carcinoma is significant as is a proliferative breast disease (confirmed in earlier biopsies), particularly when atypias were present [8, 9, 11].

Immunodeficiency disorders can result in an increased risk of malignancy. They include gene alterations. These effects constitute a predisposition for various tumor disorders frequently before age 40, yet on the whole they are rare [10].

The age of menarche has been earlier and earlier since the turn of the century as seen in Figure 2.6. It is possible that earlier exposure to endogenous hormone cycling has shifted the incidence curve to the left (younger ages) and may account for at least some of increase in breast cancer incidence

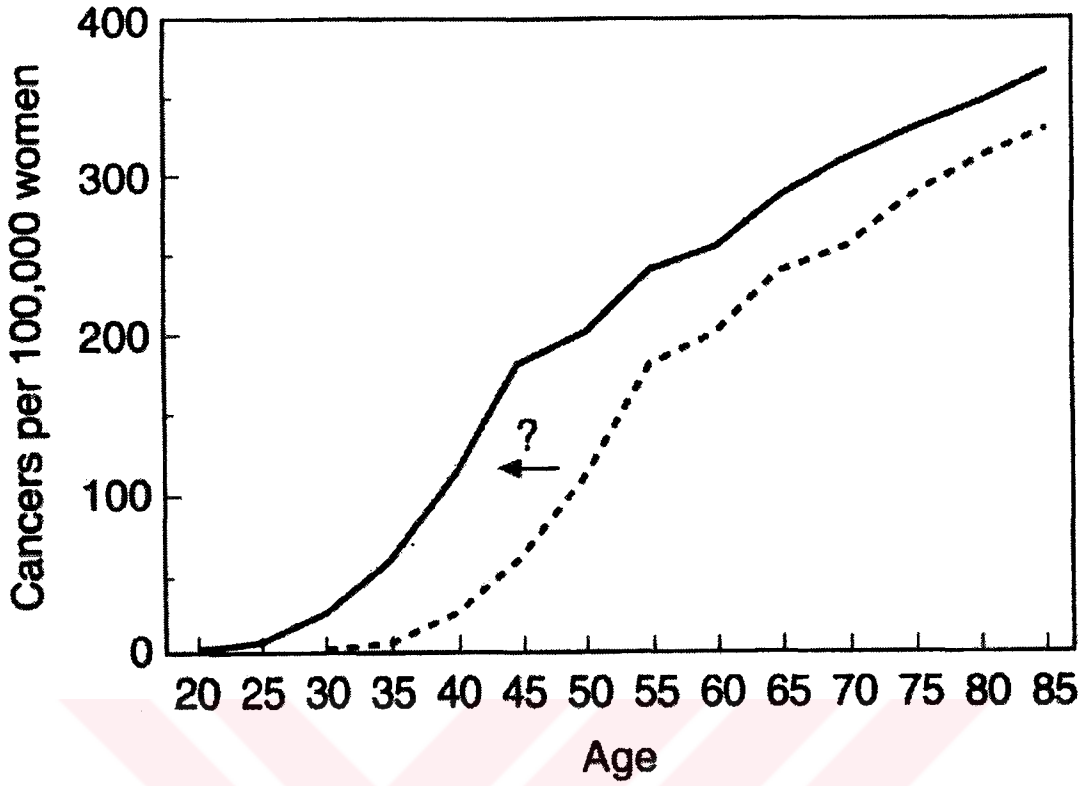

Figure 2.6 The age of menarche. This seems to get earlier and earlier.

Frequency and duration of breast-feeding has a slight impact on overall risk. In combination with other risk factors, first childbirth after the age of 30 or the absence of breast-feeding unfavorably influences the prognosis. The following medical history data may be helpful in image interpretation:

Recent pregnancy or breast-feeding can be the cause of extensive proliferation of glandular tissue, which may be misinterpreted if the physician is unaware of the patient's history. In some postmenopausal patients, hormone replacement therapy may involve extensive proliferation of glandular tissue. Newly occurring or increasing densities can be mistaken for suggestive findings if the physician is unaware of the patient's history. Published studies have described that administration of thyroid hormone can promote fibrocystic changes in the breast.

Changes after surgery or radiation therapy can produce masses, distortions or micro-calcifications that can simulate or obscure a carcinoma. The elapsed time since surgery or irradiation may also be valuable for correct image interpretation. Furthermore the following symptoms may be a hint to malignancy:

- Any changes of nipple are important. Even though deviation or inversion of the nipple can be congenital or can occur following inflammation, new development may be an important and early hint of malignancy.
- Spontaneous discharge.
- Significant aspects of any clinical findings such as skin dimpling, skin changes, palpable findings, include:
- Time when the condition was first noticed.
- Changes since the condition was first noticed.
- Results of previous examination such as surgical biopsy, core biopsy.
- If previous imaging studies exist, it may be useful to obtain these films for comparison. Comparing findings with earlier imaging studies might improve diagnostic accuracy.

2.4 Physical Examination of Breast

A complete breast examination includes the physical examination as well as a mammogram. In a screening setting, about 10% of breast cancers will only be detectable by physical examination. Initial examination of the breast involves visual inspection and palpation. When the physical examination is abnormal, subsequent diagnostic studies should always be interpreted together with clinical findings. Careful palpation is essential even with regular mammographic screening because of the following reasons:

Mammography has limited sensitivity, especially in radio-dense tissue. Approximately 10% of malignancies are only discovered because they are palpable. Palpation can detect malignant processes along the periphery of the glandular body or in the axillary tail which may escape detection at mammography. In addition to these, it is better to keep in mind that mammography does not replace careful physical examination [12].

2.5 Breast Imaging Modalities

Common breast imaging modalities include mammography, sonography and contrast enhanced magnetic resonance imaging.

2.5.1 Mammography

Mammography, has been described in details in the next chapter, is the single most important imaging method in diagnosing breast disease. After filling out a questionnaire to determine patient's risk factor, the breast is placed between two compression paddles. Since detection of small lesions in the breast will vary with breast density, the breast, is compressed to allow better visualization of the breast tissue on film. Two low dose exposures of each breast will be performed. Additional views may be needed if some areas of the breast are not clearly defined.

A radiologist, a physician who specializes in medical diagnosis by X-ray, examines the mammogram to detect any variation from the norm. If an abnormality is found, image guided biopsy and other types of diagnostic imaging such as sonography can be used to help confirm breast cancer.

There exit 3 types of cases: normal, benign and malign case. Mammographic films of each possible case are shown in Figure 2.7.a, 2.7.b and 2.7.c.

2.5.2 Sonography

Sonography, in other words breast ultrasound, is frequently used to evaluate breast abnormalities that are found with screening or diagnostic mammography or during a physician performed clinical breast exam. Under sonography and mammography examinations, a breast mass is as shown in Figure 2.8. Sonography allows significant freedom in obtaining images of the breast from almost any orientation. Its excellent accuracy in the diagnosis of cysts reduces the number of unnecessary biopsies.

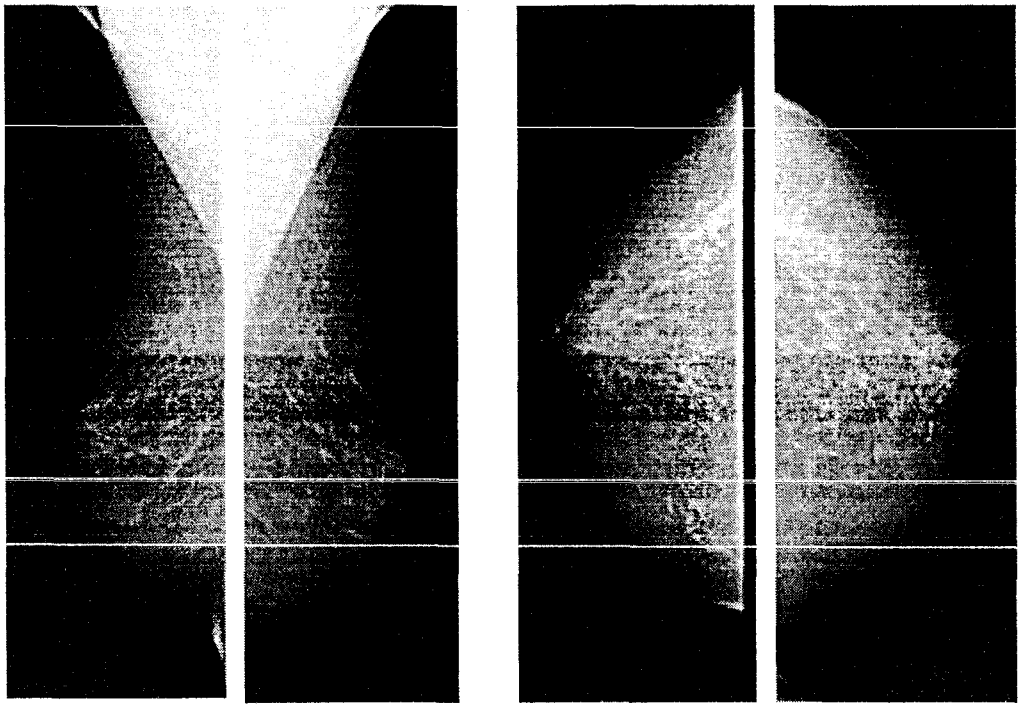

Figure 2.7.a Normal case, patient age 66, digitizer dba 21.

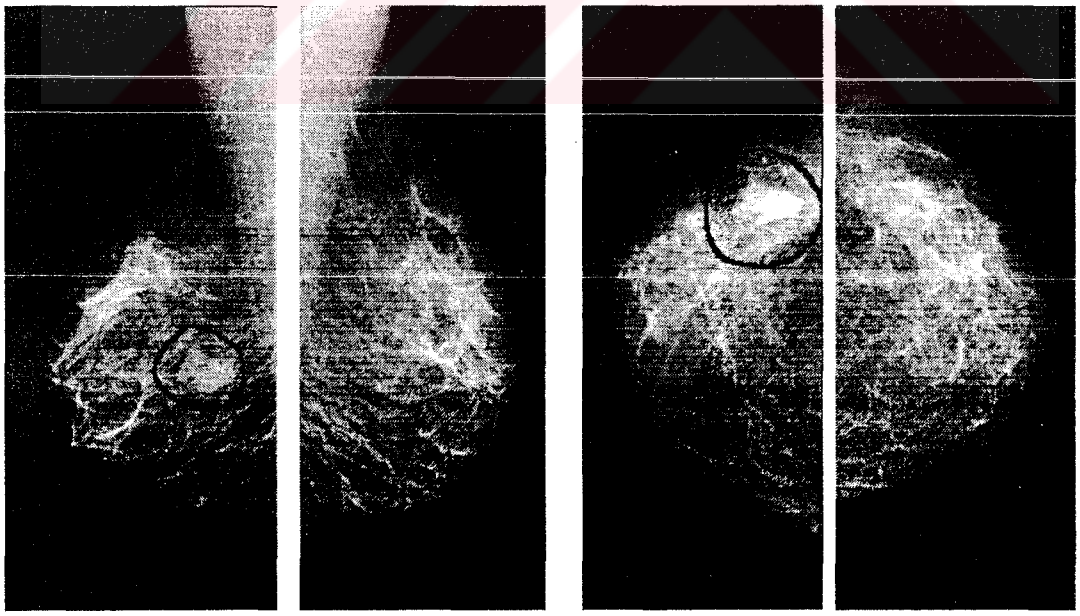

Figure 2.7.b Benign case, patient age 39, digitizer lumisys laser.

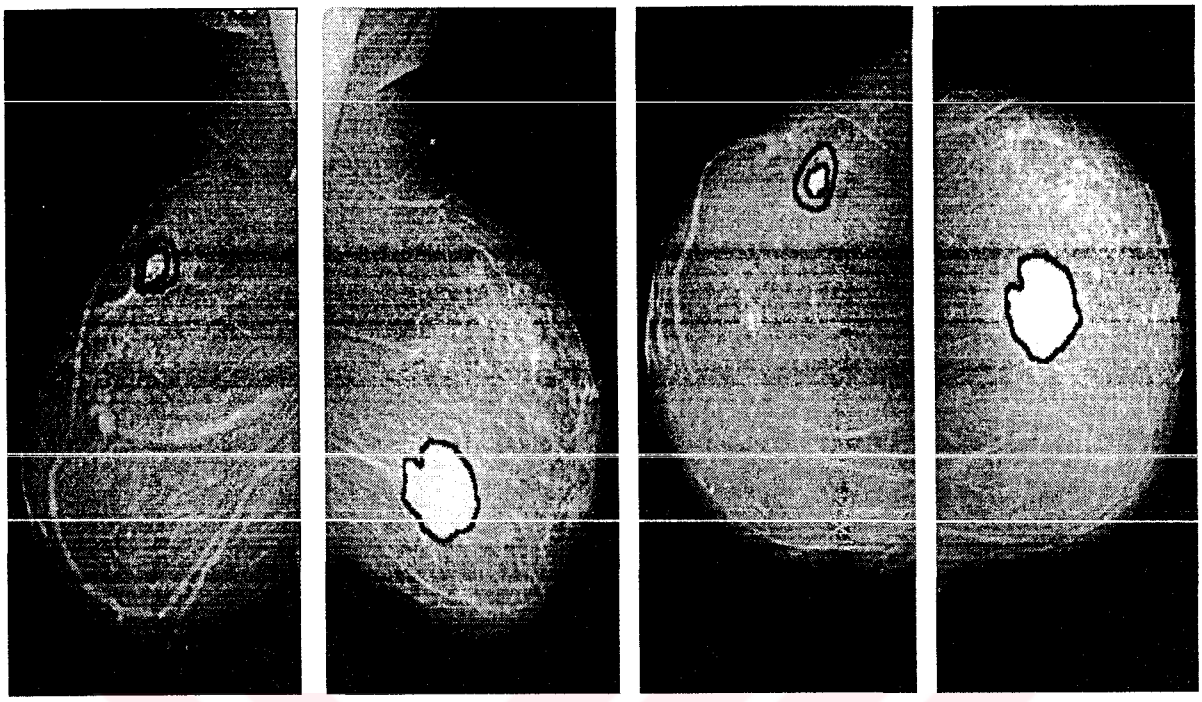

Figure 2.7.c Malign or cancer case, patient age 85, digitizer dba 21.

If a poorly demarcated solid mass in radiodense tissue can be visualized with sonography and it shows typical features of a benign lesion then sonography supports the diagnosis. If, however, any doubts exist concerning the benignity of a hypoechois lesion, biopsy is reconmmended to confirm the diagnosis.

Sonography can provide important information to confirm the presence of paplable carcinomas in radiodense tissue. The accuracy of ultrasound in detecting processes depends both on the surrounding tissue and on the lesion itself. Accuarcy in malignant processes is limited for small carcinomas and in particular preinvasie carcinoma. For this reason, sonography should never be used to rule out a malignant process without corresponding mammographic studies. Benign sonographic findings in the precense of mammographic or clinical signs of a suspected malignant process do not exclude malignancy. In view of these restrictions, the known examiner-dependent accuracy and its high physician time requirements, sonography is also not suitable as a screening modality.

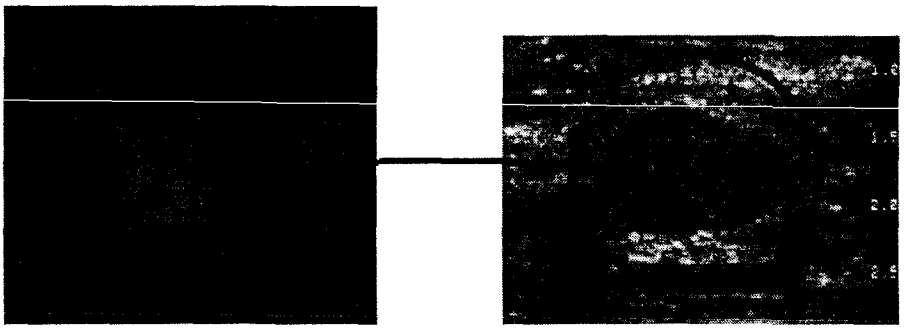

Figure 2.8 Mammogram and ultrasound view of a breast mass.

In Figure 2.9, the sonographic image of a benign breast tumor is seen while the case of ductal invasion associated with malignant breast mass is shown in Figure 2.10.

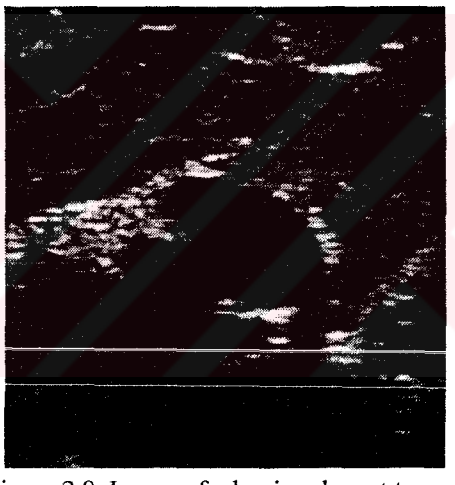

Figure 2.9 Image of a benign breast tumor.

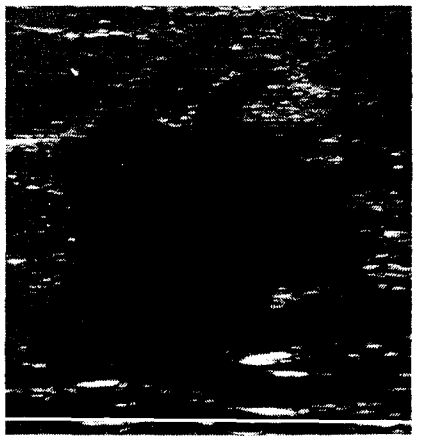

Figure 2.10 Image of a ductal invasion associated with malignant breast mass.

2.5.3 Contrast Enhanced MRI

Magnetic resonance imaging (MRI) is a relatively young method in radiology. It can be used to obtain three dimensional images of the inner parts of the human body, without using X-rays. MRI needs strong magnetic fields and uses protons, the nuclei of hydrogen atoms, to generate images. These hydrogen atoms can be found almost everywhere in the human body (about 60% of your body is water). The major advantages of MRI compared to X-ray are:

- No radiation
- The ability to generate arbitrary three dimensional views of an object of interest
- The possibility to generate images of „soft" body structures

In contrast to X-ray imaging which does not allow to distinguish tissue other than bone and flesh, MRI is ideally suited for distinction between, e.g., fatty and glandular tissue in the female breast as shown in Figure 2.11. In this coronary section, the fat is very bright due to its high percentage of hydrogen, whereas the glandular tissue, which does not contain as much hydrogen, shows only average brightness.

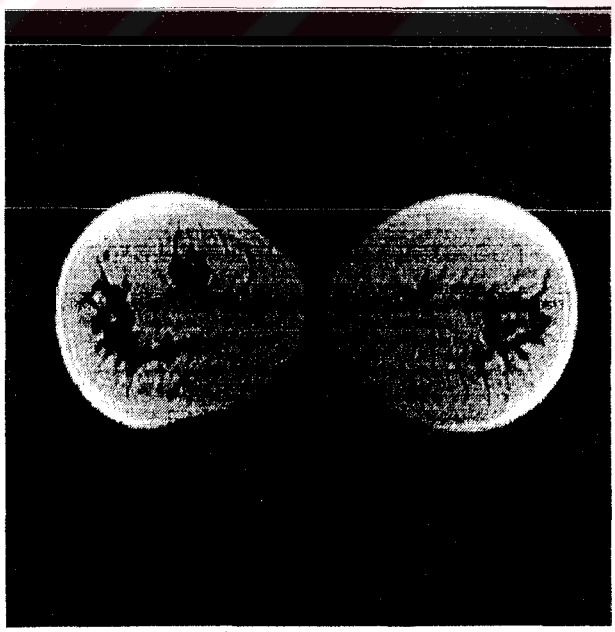

Figure 2.11 A 2-D section of an MRI image of a female breast.

A rather new, and very promising technique to investigate the female breast for cancer is Magnetic Resonance Imaging (MR-mammography) using a contrast medium, usually Gadolinium-diethylene-triamine-pentaacetic acid (Gd-DTPA). Compared to mammography using X-rays this method has the following advantages:

- No gamma-ray exposure.
- Acquisition of three-dimensional images.
- Additional information for difficult cases.

The method is based on the fact that a carcinoma is usually well vascularized due to its strong growth. Thus a contrast medium will quickly accumulate in the tumor. Using a dedicated MRI-coil, one three-dimensional image of the breast is taken before application of the contrast medium and several images are taken afterwards, about one per minute. In this process a large number of two-dimensional images is achieved, somewhere between 50 and 400, that have to be examined and evaluated by the radiologist.

The evaluation essentially consists of finding suspicious regions in the images, calculating the absorption of contrast medium in those regions, and deriving the diagnosis from that data. This process is obviously a very time-consuming task. Thus, it cannot be used as a screening method. Many publications show that MR-mammography has a very high sensitivity. In addition to this it is an expensive examination method and to avoid too many false positive calls in asymtomatic patients, MR-mammography should predominantly be used in patients with a significantly increased risk of malignancy or a suspicion of malignancy, if evaluation by conventional imaging is limited.

In Figure 2.12.a, MR breast image of a patient is shown. In the image a benign mass is labeled as (A) while invasive lobular carcinoma is labeled as(B). The image is taken before contrast injection and then after contrast injection as shown in Figure 2.12.b. The final picture, as seen in Figure 2.12.c, shows the post-contrast image subtracted from the pre-contrast image in order to display regions that are highlighted by the contrast. In particular, the subtracted image highlights the speculated area that shows invasive lobular carcinoma.

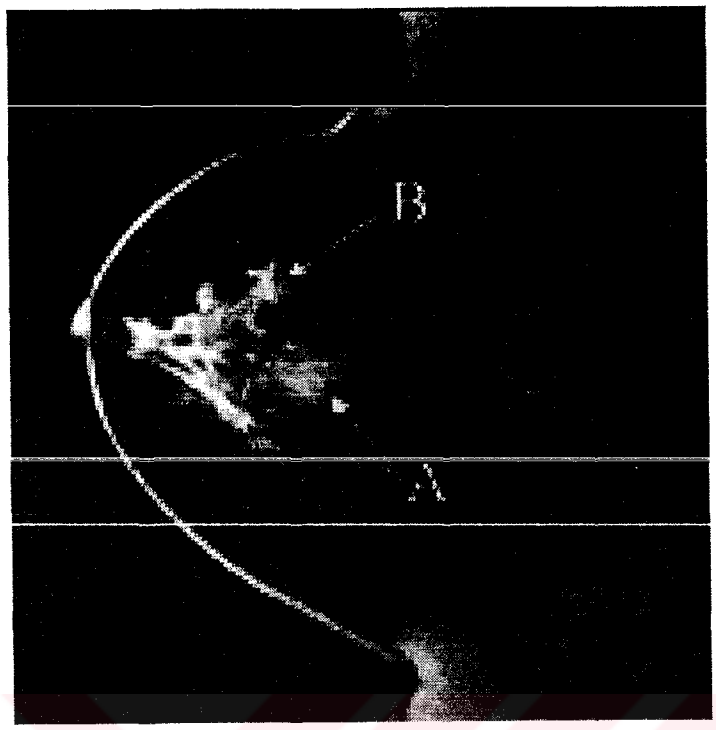

Figure 2.12.a MR-Mammogram before contrast injection.

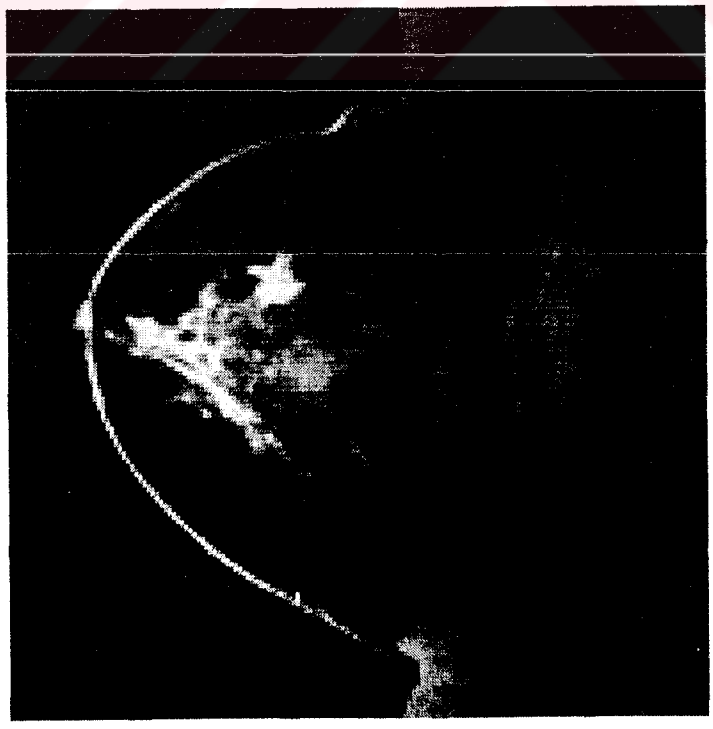

Figure 2.12.b MR-Mammogram after contrast injection.

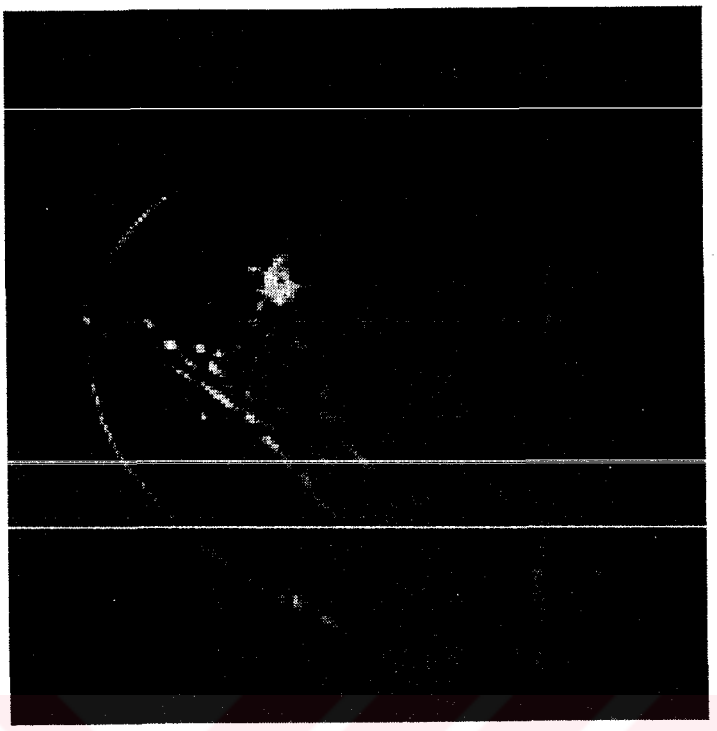

Figure 2.12.c Subtracted image, only the cancer is visible.

3. MAMMOGRAPHY

3.1 Introduction

Mammography is the single most important method in diagnosing breast disease. Its areas of application include screening and problem solving. Mammography is the only imaging method to date that is suitable for screening. On the other hand, aside from a few exceptions, such as very young patients, mammography is always indicated as a diagnostic method in symptomatic patients.

Conventional mammography requires that all components of the imaging sequence be performed properly. This includes the X-ray device and its components including the X-ray tube and its cathode, anode, focal spot, window, filtration, collimation, source to image distance (SID) and the compression system automatic exposure control (AEC), the detector and its components, including the cassette, film and screen or the solid state detector, the film processor and its components, the interpretation system (Figure 3.1). All of these components are interrelated. Alterations at any point in the sequence can influence the operation of the other elements. Even though the detector and the display are different between conventional mammography and digital mammography, the interactions of the constituents of the system must be carefully matched and properly adjusted (Figures 3.2 and 3.3).

3.2 Components of Mammographic Imaging

3.2.1 X-Ray Tube

Mammography requires special powerful tubes that produce particularly low energy radiation in comparison to other diagnostic X-ray tubes to achieve the required high tissue contrast. This is necessitates the use of special targets and filters. Since the radiation needed originates in a small focal spot, the exposure time should be as short as possible to avoid motion blurring.

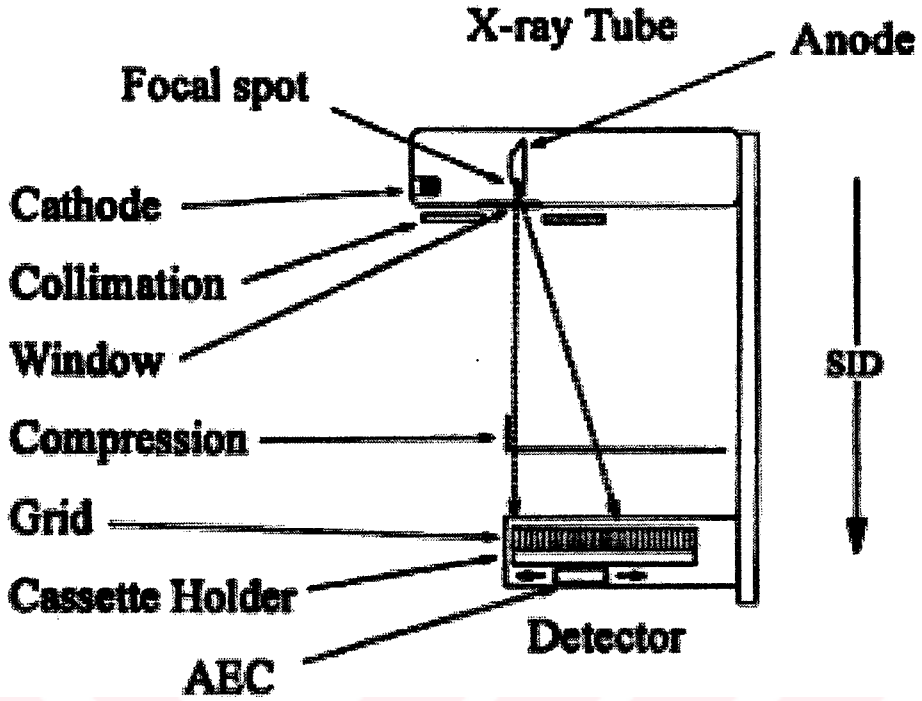

Figure 3.1. The components of X-ray device.

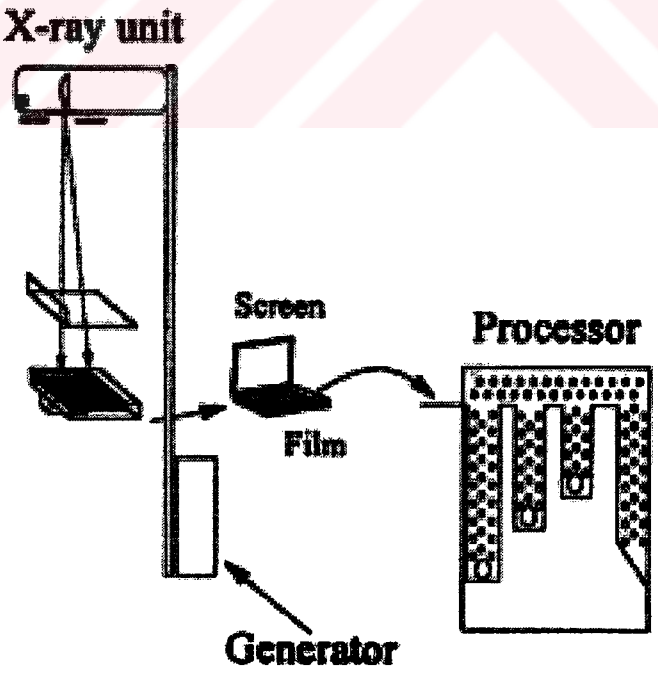

Figure 3.2 Conventional mammography.

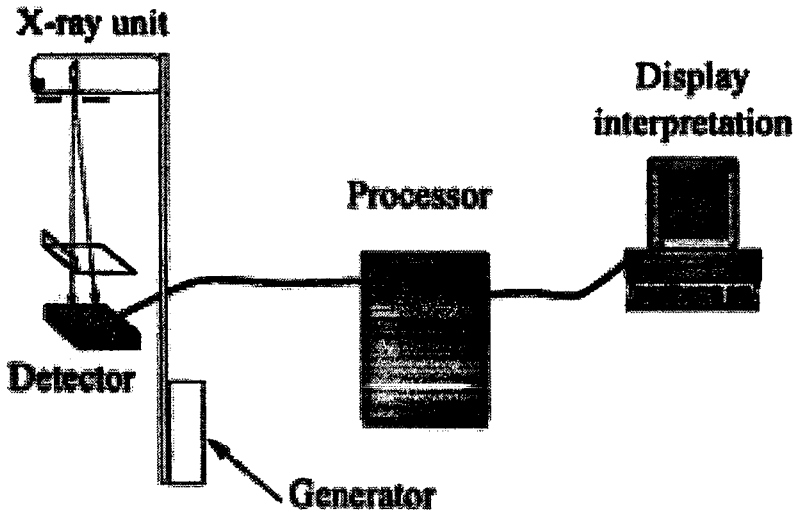

Figure 3.3 Digital mammography.

3.2.2 Sharpness

To achieve the required spatial resolution, mammography tubes must have an extremely small focal spot. A nominal focal spot size smaller than 0.4 is the current practice. However a nominal focal spot size of 0.4 means that the diameter in each direction will be between 0.4 mm and 0.6 mm. The local projection of the width of the focal spot will vary according to its distance from the chest wall and the angulation of the tube.

In addition to minimal focal spot size, the proper geometric configuration of the focal spot, object and the image receptor is important in achieving the necessary sharpness. Use of a small focal spot, the shortest possible distance between the object and the film and the longest possible distance between the focus and the film will minimize geometric blurring.

3.2.3 Radiation Spectrum

The radiation spectrum produced in X-ray tubes is not monoenergetic but consist of a spectrum of radiation energies. This spectrum comprises X-ray bremsstrahlung (braking)

and the characteristic radiation determined by the target material. Since the spectrum of imaging radiation greatly influences contrast and radiation dose, the following physical aspect should be considered:

- With low-energy radiation, slight differences in the radiodensity of soft tissue of the breast that would otherwise remain undetected can be visualized with high contrast. Increasing energy of the radiation decreases soft tissue contrast. Also the radiation spectrum must have sufficiently high energy for adequate penetration of thick breasts and breasts with abundant fibrotic of glandular tissue.

- Radiation with insufficient energy will not penetrate the breast even with long exposure time. Such radiation is not suitable for imaging at all. It will unnecessarily increase the radiation dose and since dense tissue cannot be penetrated it will produce an inadequate image. Thus, higher-energy radiation is required in dense breasts and in thick breasts. With the optimum radiation energy selected, the *absorption* is higher in radiodense tissue fibrotic tissue, glandular tissue, and malignant tissue) than in radiolucent tissue (fat or loose connective tissue). These differences in absorption produce the image pattern.

Since too large a component of high energy reduces the contrast and too high a component of low energy results in excessive radiation exposure, it is advisable to adapt the radiation spectrum as closely as possible to the thickness and density of the breast. The radiation spectrum is determined by the following factors:

- The target filter combination of the X-ray tube.
- The peak kilovoltage (kVp) setting on the X-ray unit.

3.2.4 Target / Filter Combination

The radiation spectrum created at the target depends on the kVp setting and on the target material. The radiation spectrum of molybdenum targets contains a higher proportion of low-energy radiation (including characteristic peaks at 17.5 and 19.6 keV)

than do the spectra of tungsten or rhodium tubes. Selective filtering is used to adapt the radiation spectrum of a given target as closely as possible to the specific requirements.

Selective filtering can be used to suppress the low-energy components of the spectrum that would represent unnecessary radiation exposure because they are absorbed in the breast. It will also reduce the energy components above the K absorption edge characteristic of the selected filter material, essentially permitting a narrow spectral range directly below the K absorption edge to pass. Any filter is particularly efficient at absorbing that part of the radiation whose energy exceeds a limit, referred to as the K absorption edge, specific to the filter material.

The effective spectral range can thus be defined by selecting the target and filter material and the thickness of the filter. Commercially available target/filter combinations include molybdenum/molybdenum molybdenum/rhodium, rhodium/rhodium or tungsten/ molybdenum, and tungsten/rhodium.

The radiation quality from a molybdenum/ molybdenum or tungsten/molybdenum target/filter combination is suitable for most breasts. The combinations tungsten/molybdenum molybdenum/rhodium, tungsten/rhodium, and rhodium/rhodium provide, in this order, increasingly high-energy radiation spectra. They permit better penetration of large and mastopathic breasts with abundant glandular, fibrotic, and connective tissue, resulting in higher image quality and a reduction and unnecessary radiation exposure.

3.2.5 Peak Kilovoltage (kVp)

A higher kVp setting increases the relative proportion of high-energy radiation in the respective spectrum, whereas a lower kVp setting increases the relative proportion of low-energy radiation. Selecting the proper kVp setting, target material and filter material according to breast thickness and density is important. Since the optimum kVp for a target/filter combination is not applicable to every subject, automatic exposure control systems are provided to make it easier to match kVp to breast thickness and density. Depending on the manufacturer, these systems may select or suggest the proper settings.

3.2.6 Penetration: The Heel Effect

The heel effect of the X-ray tube is also exploited to compensate for varying penetration in the chest wall and nipple. The heel effect means that the intensity of rays emitted by the target is not uniform throughout the beam [13].

More of the rays that leave the target at obtuse angle will be absorbed by the target than those leaving at acute angle due to the longer path they have to travel in the target. Since the thickness of the breast is greater close to the chest wall than near the nipple, it is best when the area of maximum radiation intensity lies near the chest wall. This is achieved by positioning the target opposite the cathode, which is closer to the chest wall. The intensity distribution of the radiation can be influenced by slightly angling the X-ray tube. However, this alters the projection of the focal spot.

3.2.7 Scattered Radiation

In every radiograph of the breast, scattered radiation is produced in the tissue. In denser and thicker glandular tissue, more scattered radiation occurs than in the thinner, fatty, transparent tissue. Increasing amounts of scattered radiation result in progressive loss of contrast.

3.2.8 Scattered Reduction

The grid is placed between the breast and the image receptor for a screen-film system to reduce undesired scattered radiation that impairs image quality. Grids consist of strips of lead that absorb obliquely oriented radiation whereas radiation parallel to the lead strips passes through. The lead strips are focused on the focal spot. During the exposure, the grid rapidly moves perpendicular to the path of the beam and to the orientation of the strips to prevent the strips from appearing on the mammogram as thin lines that mar the image.

The efficiency of the grid depends on the height of the strips and the strip spacing. The ratio of strip height to strip spacing is known as the grid ratio. The larger the grid ratio, the greater the efficiency of the grid but also the greater the required radiation dose. For this reason, only grid ratios of 4:27 or 5:30 are recommended for mammography [14].

Since the grid absorbs both scattered radiation and a small proportion of useful radiation, it requires a longer exposure time and therefore an increased radiation dose. Exposures with a scattered radiation grid require a grid exposure factor of approximately 2.5. The use of more sensitive screen-film systems has compensated for this increased dose, compared with earlier grid-less mammographic techniques.

Significant increase in image quality fully justifies the increased radiation dose required by the grid and grid mammography has superseded griddles mammography. Griddles mammography can only be per-formed without significant loss of quality in very small, compressed, and fatty breasts in the interest of reducing radiation exposure.

The second important method of reducing scattered radiation consists of sufficient compression of the breast. By reducing breast thickness, compression reduces the proportion of scattered radiation, thus reducing the dose and improving the image contrast [15]. Other options for reducing scattered radiation include air-gap technique. The air gap, which is effective only in conjunction with good collimation, is used for scatter reduction in magnification mammography.

3.3 Accuracy

A basic knowledge of the accuracy of mammography is an important prerequisite to properly judge its value in screening and clinical use.

3.3.1 Sensitivity

Realistically, mammography has a sensitivity of about 90%, about 10% of all carcinomas are not detected initially by mammography. When mammographic screening is performed, about 25-35% of the carcinomas become apparent between screening examinations, usually by manifesting clinical symptoms. They are called interval carcinomas. It is important to know that numerous carcinomas detected at screening are retrospectively visible on the previous examination, mostly as some uncharacteristic change [16, 17, 18].

Thus mammography does not provide 100% sensitivity. There exists a threshold for mammographic detection of malignancy, which depends on tumor size, tumor type, and surrounding tissue. These limitations must the kept in mind, particularly for diagnostic mammography. For screening, however, mammography is the only method that allows reproducible and reliable detection of a prognostically relevant number of nonpalpable carcinomas, at an acceptable rate of false positive calls and at acceptable expense.

Overall sensitivity of mammography in fatty tissue is excellent. It decreases as radiodensity increases. This means that mammography has a lower sensitivity in radiodense tissue and therefore a negative mammogram does not eliminate the need for further workgroup of otherwise indeterminate or suggestive palpable findings in dense tissue [19, 20].

Mammography is highly sensitive in detecting carcinomas containing microcalcifications, and this sensitivity is largely uninfluenced by the radiodensity of the surrounding tissue. Carcinomas account for about 50% of all cancers, including approximately 30-40% of all invasive carcinomas and about 90% of carcinomas in situ currently detected. Since these are generally not palpable but have excellent cure rates, mammography plays a decisive role in early detection.

3.3.2 Specificity

Mammography is specific in only a few cases:

- Absence of malignancy can be diagnosed reliably in fatty breasts (provided the area in question is included on the mammogram).
- A definitive diagnosis of a benign lesion is possible for a typical oil cyst, a hamartoma, a lipoma, a typically calcified fibro-adenoma or lymph nodes with typical mammographic features.
- A quite reliable diagnosis of a benign tumor or cyst (> 98% correct) is possible in the case of a typical well-circumscribed mass,

In the majority of clinically or mammographically detected changes, mammography, however, is nonspecific and only permits likelihood statements [21, 22].

The specificity of the diagnosis of a carcinoma is quite high for spiculated masses, as well as for pleomorphic and cast-like microcalcifications with ductal distribution. However; a spiculated mass can also be caused by an area of fist necrosis or a radial scar. The size of the findings decisively influences the expected specificity of the mammographic study. In fact most nonpalpable carcinomas, in particular small carcinomas, appear as nonspecific changes. Unless the examiner is only looking for large, obvious findings, one has to be aware that only 1 of every 5 to 10 mammographically suspicious changes will correspond to malignancy.

Further diagnostic studies, including additional views, sonography and percutaneous biopsy can improve this rate so that more than half of the excisional biopsies of nonpalpable abnormalities will be performed for a malignancy [23].

3.4 Compression and Positioning

Mammographic positioning involves obtaining the best compression and correct positioning to ensure that the entire glandular body is imaged.

3.4.1 Compression

Adequate breast compression is one of the most important prerequisites for obtaining high-quality mammograms with the best possible visualization of pathologic changes [24]. The contributions of compression to mammography quality are as follows:

- Good compression improves the resolution by reducing the distance between the image receptor and objects, which are further from the image receptor. This reduces geometric blurring.
- It reduces motion blurring.
- It improves contrast since the reduced thickness of the breast significantly reduces scattered radiation.

- It improves contrast since low-energy radiation, which provides higher image contrast, can be utilized in penetrating tissue of reduced thickness.

- It permits higher contrast in the area of interest since by equalizing the thickness of the tissue; it reduces the necessary object range.

- It permits visualization of small areas of pathology buried in the glandular tissue since normal tissue can be spread, whereas malignant foci will persist due to their firmer consistency.

- It permits a significant reduction in the dose by reducing the thickness of tissue to be penetrated.

These advantages illustrate the importance of achieving the best possible compression. As an important prerequisite for high image quality, early cancer detection may depend on it. Naturally, the best possible breast compression can only be achieved with the patient's cooperation and must never be obtained against her will. Despite every effort to achieve optimum image quality, the technical staff must appreciate that the sensitivity of the glandular body to pressure varies. Patients differ in their willingness to endure pain or discomfort for good diagnostic results.

It is thus essential to briefly discuss the need for compression with the patient and to obtain her understanding, cooperation, and motivation. For this reason, the patient should be informed that some of the smallest and earliest cancers can only be visualized with compression and that compression significantly reduces the dose of radiation. She should be told that there is no way in which compression can cause a carcinoma, which is a common concern.

Since compression of tissue with low interstitial water content is less painful and the density of the breast decreases and image quality improves as the water content decreases, the mammographic examination may be more comfortable for women during the first half of the menstrual cycle. When compressing the breast, it is important that all of the glandular tissue is spread as evenly as possible and that no skin folds are present. Compression of unevenly distributed glandular tissue is more painful and the folds may cause densities that interfere with the diagnosis.

3.4.2 Positioning for Standard Views

Except in unusual instances, all mammographic examinations should be obtained in two planes. So-called single-view mammography, which essentially consists of only the mediolateral oblique view, has been used for screening due to the lower cost involved. Because of its reduced sensitivity and specificity, most mammography experts do not regard single-view mammography as sufficient for diagnostic purposes or cost-effective. Many patients require repeated examinations. For this reason, mammography in only one plane should be reserved for exceptional cases.

Whereas standard views permit reliable identification or exclusion of malignant processes in most patients, additional views should be used liberally whenever mammograms in standard imaging planes are inconclusive or do not visualize the findings completely. Any additional view is preferable to an unnecessary biopsy or a carcinoma that goes undetected. The mediolateral oblique and the craniocaudal view is used in combination have become the international standard views [25].

3.4.3 Mediolateral Oblique View

The mediolateral oblique (MLO) view is regarded as the most important view since it best visualizes the tissue adjacent to the chest wall and the axillary tail. It is the view that is most likely to include all the breast tissue. It is designed to maximize visualization of the lower axilla and the upper outer quadrant. Most carcinomas can be visualized in the mediolateral oblique view. A well-positioned mediolateral oblique projection and the obtained mammogram are shown in Figures 3.4.a and 3.4.b.

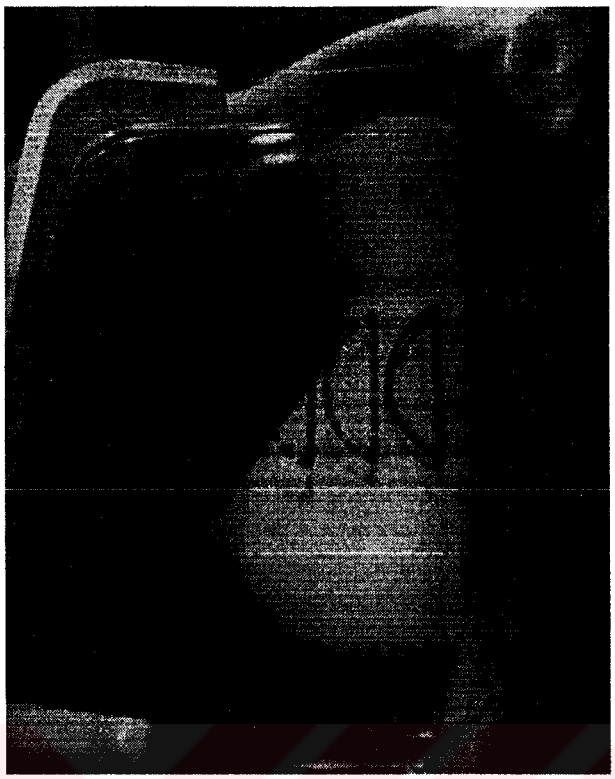

Figure 3.4.a A well-positioned mediolateral oblique projection

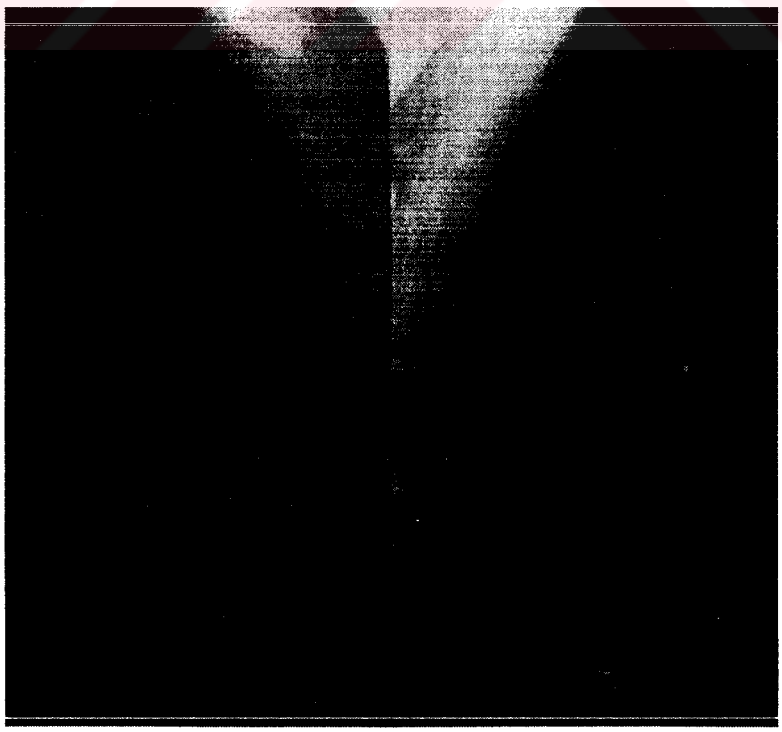

Figure 3.4b Mammogram.

3.4.4 Craniocaudal View

The mediolateral oblique view is routinely supplemented by the craniocaudal view in which the beam travels from superior to inferior. During the projection to prevent the opposite breast from pushing the patient away from the detector, the breast is placed upon on the detector as seen in Figure 3.5.

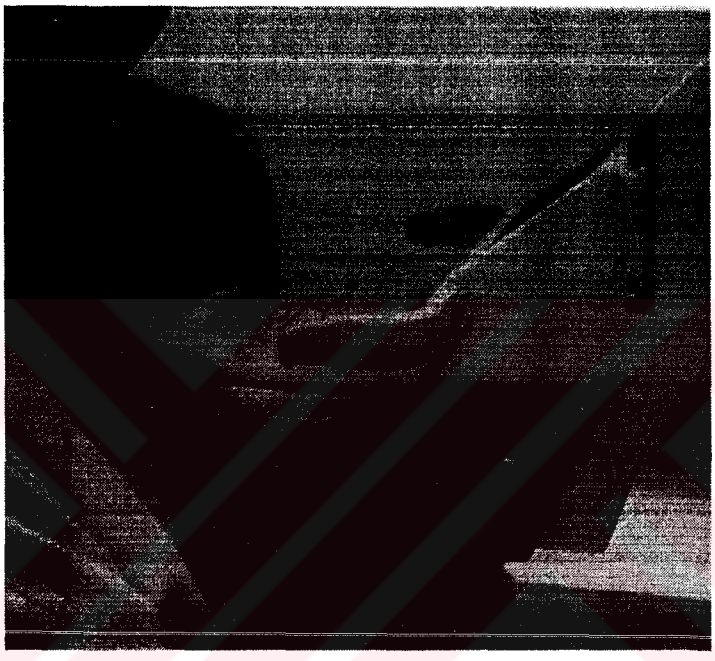

Figure 3.5 A craniocaudal projection.

3.5 ACR Standard for the Performance of Diagnostic Mammography

The goal of all mammography is to foster detection of unsuspected breast cancer in either breast, but diagnostic mammography is intended to provide specific analytic evaluation of patients with clinically detected or screening-detected abnormalities. Diagnostic mammography should lead to conclusions about the patient's clinical symptoms or mammographically detected findings to enable specific management recommendations. To achieve this consultative function, additional separate studies such as sonography, magnetic resonance imaging, may be indicated.

The request for diagnostic mammography should be regarded as a request for diagnostic consultation. In addition to routine mammographic evaluation, a diagnostic mammogram may include additional views more then mediolateral oblique and craniocaudal projections.

Additional separate studies or procedures such as sonography, fine-needle aspiration, core needle biopsy, MRI and others may be indicated to complete the diagnostic assessment. Diagnostic mammography should be performed under the direct supervision of an interpreting physician qualified in mammography. The patient's history, symptoms and signs, reported findings on physical examination, and results of prior mammography, if performed, will focus the diagnostic breast evaluation.

The goal of diagnostic mammography is to provide additional information about patients who have signs and/or symptoms of breast disease; radiographic findings of concern, or those conditions where the interpreting physician deems direct supervision appropriate. Information gained from this study should lead to specific interpretive conclusions and/or further diagnostic and management recommendations. Indications for diagnostic mammography include:

- Specific focus of clinical concern including, but not limited to, mass, indurations, axillary lymphadenopathy, some types of nipple discharge, skin changes, or persistent or focal areas of pain or tenderness.
- Possible radiographic abnormalities detected on screening mammography.
- Short interval follow-up (e.g., less than 1 year) for clinical or radiographic concerns.
- Any patient whose examination requires direct involvement of the radiologist for special views, breast physical examination, or consultation.
- Women who have been treated for breast cancer.

Although the potential risk of mammography to the fetus is negligible and mammography is not contraindicated for the pregnant patient, if the patient is known to be

pregnant, the suggested procedure in the paragraph below regarding pregnancy should be followed.

All imaging facilities should have policies and procedures to reasonably attempt to identify pregnant patients prior to the performance of any diagnostic examination involving ionizing radiation. If the patient is known to be pregnant, the potential radiation risks to the fetus and clinical benefits of the procedure should be considered before proceeding with the study. Statement of clinical concerns or indications should be obtained at the time patients are scheduled for diagnostic mammography. Palpable concerns, foci of pain, or point tenderness may be marked with small radiopaque markers. Thin wires may be used similarly to mark biopsy scars. Radiographic demonstration of surface markers may provide positioning guidance for routine; spot compression, tangential, and other views over the image receptor near the axilla for identification of laterality and view.

A diagnostic mammogram may include additional views to evaluate an area of clinical or radiographic concern. Additional mammographic views might include spot compression, spot compression with magnification, tangential views, or other special views. When selecting a view, the proximity of the area of concern to the image receptor should be considered.

Evaluation of the augmented breast should include standard craniocaudal (CC) and mediolateral oblique (MLO) as well as implant-displacement views, if possible. Evaluation of the augmented breast may require, in some cases, additional tailored or special views for optimal visualization of breast tissues. Further evaluation of implant integrity may include the separate use of ultrasound or MRI.

Adequate documentation of pertinent patient and technical information is essential for high-quality patient care. Image labeling should include a permanent identification label that contains:

- The facility name and location
- Examination date
- Patient's first and last names
- Identification number and/or date of birth

- Technologist's identification number, initials, or other symbol
- Cassette (screen) and/or receptor number
- Mammographic technical parameters such as target material, filtration material, kVp, mAs, and compressed thickness (cm) of the breast may also be included.

Retention of mammographic images should be consistent with clinical need and be in compliance with federal and state regulations and local health care facility requirements and regulations. The image also should include a radiopaque marker placed over the image receptor near the axilla for identification of laterality and view. In addition to the specifications stated for screening mammographic units discussed in ACR Standard for the Performance of Screening Mammography section, equipment used for diagnostic mammography must have magnification and spot-compression capability. Comparison with prior breast imaging studies may be an important part of diagnostic mammography. If previous breast imaging studies are needed for assessment of mammographic findings, an attempt should be made to obtain them.

All reports in the categories of suspicious abnormality or highly suggestive of malignancy should be directly communicated to the referring physician, other health care provider, or an appropriate representative who will be providing clinical follow-up. The actual or attempted direct communication should be documented. The clinical or radiographic concerns that prompted the diagnostic mammogram should be acknowledged.

The location of mammographic abnormalities can be indicated by using clock position; distance from the nipple; quadrant of the breast; and/or location within the anterior, middle, or posterior third of the breast. The diagnostic mammogram report should describe pertinent observations, establish levels of suspicion of malignancy based on the imaging findings, and provide specific recommendations for patient diagnosis and management.

If additional, separate breast imaging studies or procedures are performed or are available, they may be correlated in the diagnostic mammography report. Separate studies may be listed individually or combined as breast imaging in heading of report. The ACR Breast Imaging Reporting and Data System (BI-RADS™) is available to provide a

framework for reporting, lesion assessment, imaging-pathologic correlation, and quality improvement. Accurate record keeping, patient tracking, and outcome analysis are important for effective, diagnostic mammographic imaging evaluations.

3.6 ACR Standard For The Performance Of Screening Mammography

Periodic mammography screening of asymptomatic women has been shown to reduce breast cancer mortality. The principles of quality for mammography do not differ basically from those applicable to other radiological examinations. Key points to be considered are the criteria for credentialing professionals, equipment specifications, monitoring and maintenance schedules, standards for image quality, standardized image evaluation procedures, meticulous record keeping, and periodic review of data for outcomes of the mammography services.

Screening mammography is a radiological examination to detect unsuspected breast cancer in asymptomatic women. This examination may be performed without a physician in attendance. Screening mammography is indicated in asymptomatic women at least 40 years of age. It is reasonable to institute screening mammography at an earlier age in women with high risk factors. Some women may not be candidates for screening mammography

Asymptomatic women at least 40 years of age should have an annual mammographic screening examination. It is unclear at what age, if any; women cease to benefit from screening mammography. Because this age is likely to vary with the individual depending on her overall health, the decision as to when to stop routine mammography screening should be made on an individual basis by each woman and her physician.

For screening mammography, the term "self-referred" is defined as a woman who refers herself for medical services and who does not have an identified referring physician or other healthcare provider. To maximize utilization of screening, direct access by individuals is permissible without requiring physician referral in advance. However, screening facilities that elect to accept such patients must have procedures for referral to a qualified physician who has agreed to assume clinical responsibility.

All imaging facilities should have policies and procedures to reasonably attempt to identify pregnant patients prior to the performance of any diagnostic examination involving ionizing radiation. If the patient is known to be pregnant, the potential radiation risks to the fetus and clinical benefits of the procedure should be considered before proceeding with the study.

In the U.S.A., interpreting physicians, medical physicists, and radiological technologists who work in mammography must meet the Mammography Quality Standards Act (MQSA) final rule as published by the Food and Drug Administration (FDA). Mammography equipment must meet MQSA final rule as published by the FDA. Unfortunately, there are not any similar regulations as yet in our country.

The examination should ordinarily be limited to craniocaudal and mediolateral oblique views of each breast. On occasion, supplemental views may be required to visualize breast tissue completely or optimally, but such views are not ordinarily part of the routine screening examination. When pathology is suspected, a recommendation for additional imaging studies, diagnostic mammography, or biopsy may be warranted. If a breast physical examination is not available at the screening site, women should be informed that physical examination is a complementary and necessary procedure. An attempt should be made to obtain prior mammograms when the interpreting physician deems it necessary. Under the MQSA final rule, facilities must provide original films.

4. MAMMOGRAPHIC MASS SHAPE ANALYSIS

4.1 Introduction

Approximately 80 to 85% of localized breast cancers are diagnosed by the mammographic appearances of the tumor. The primary features that indicate malignancy are related to the tumor's density, size, shape and borders. The malignant mass often infiltrates the surrounding tissue. As a result, the boundary may be compromised of fine linear strands extending irregularly outward from the central mass.

Because of the malignant mass pathology, the shape of the mass can be used to discriminate between malignant and benign masses. In Figure 4.1, the morphological spectrum of breast masses frequently seen on mammograms is illustrated [26].

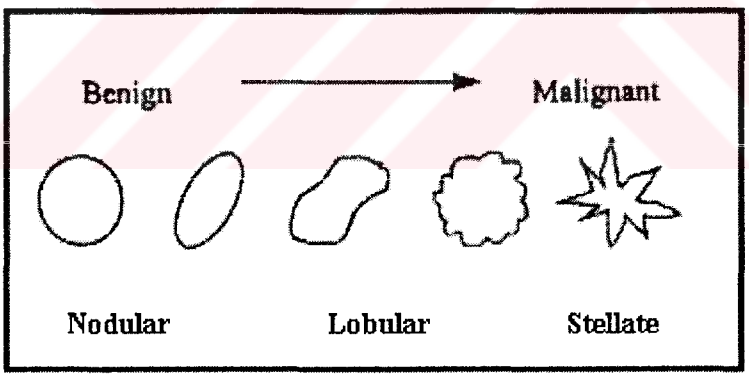

Figure 4.1 Morphologic spectrum of mammographic masses.

In Figures 4.2.a-4.2.d, several mammographic mass shapes are shown with respect to the types of margins and biopsy proven classifications. These mass shapes are taken from the mammographic mass database developed during this study.

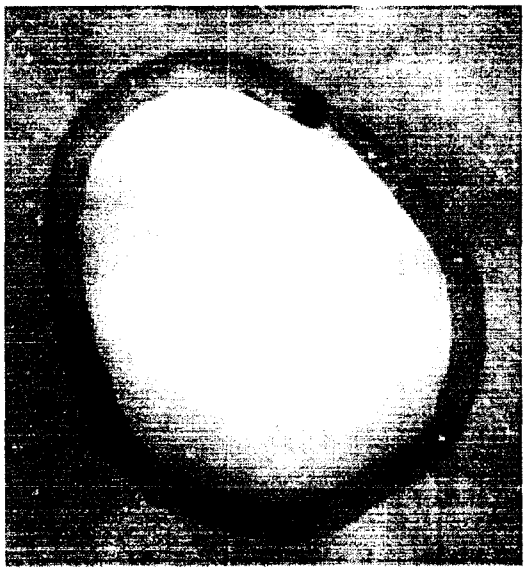

Figure 4.2.a A nodular benign tumor.

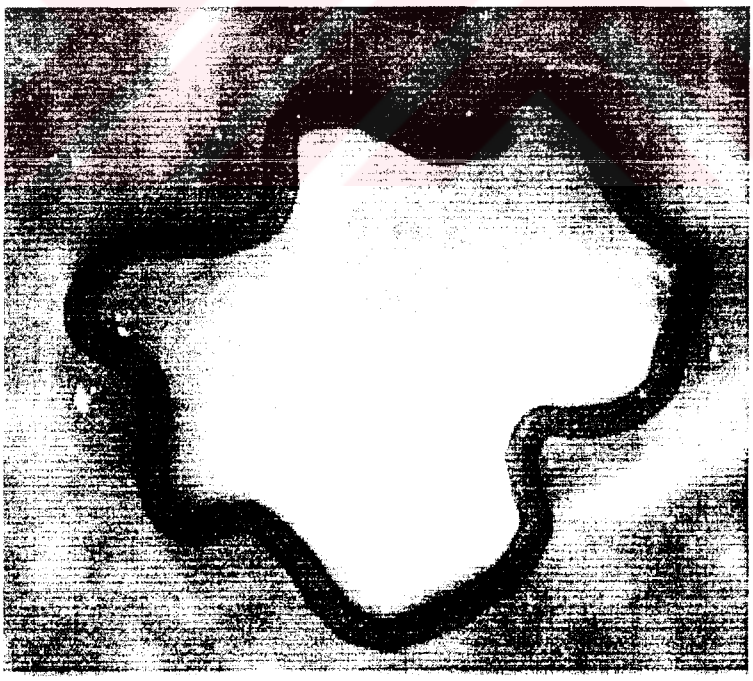

Figure 4.2.b A lobular benign tumor.

Figure 4.2.c A lobular malign tumor.

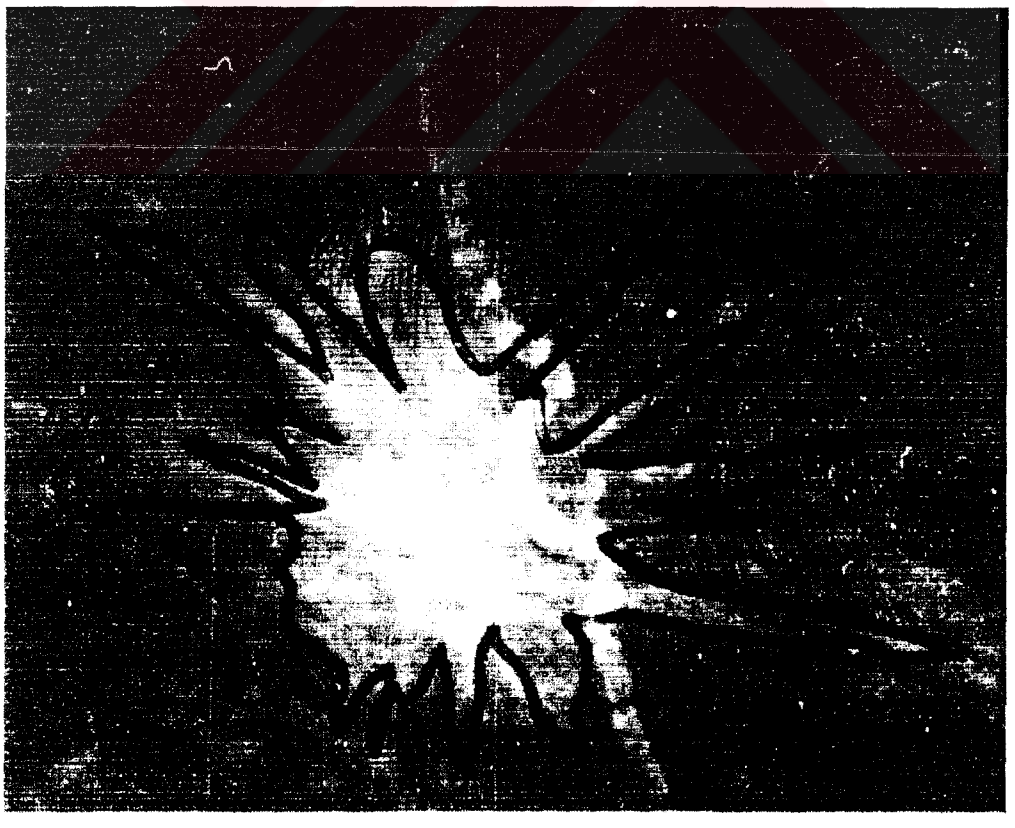

Figure 4.2.d A stellate malign tumor.

As described by Gallager and Martin, the shapes of cancerous masses often fall into two classes: lobular or stellate. They found that the outline, the margins and the shape of the mammographic masses could be used in the classification of the masses as either benign or malignant [27].

In the present thesis work a computer software has been developed to automatically compute the shape parameters of pre-segmented mammographic images. The software, after detecting mammographic mass shape from pre-segmented images, computes geometric parameters, moment based shape features and Fourier descriptors. The parameter values are stored in an excel file, *Parameters.xls*, given in Appendix A. Details of the algorithms used in the computations are presented in the next sections

4.2 Simple Geometric Parameters

Several qualitative and quantitative techniques have been developed for characterizing the shape of masses in an image. These techniques are useful for classifying masses in a pattern recognition system and for symbolically describing masses in an image understanding system. Some of the techniques apply only to binary valued images while others can be extended to gray level images [28]. In the present work, the simple geometric parameters of a mammographic mass shape computed include area, centroid, perimeter, normalized circularity, radial distance mean, radial distance standard deviation and area ratio.

4.2.1 Area

The most trivial shape parameter is the area of a mass. In a digital binary image the number of pixels that belong to the mass gives the area. In the matrix or pixel list representation of the mass, area computing simply means counting the number of pixels. On the other hand, chain code of the mass can also be used to calculate the area.

4.2.2 Centroid

The centroid, also called as the center of gravity, is the balance point of a binary image. Detecting the centroid of a mass is the first step during radial distance

measurement. Centorid of an image (x_c, y_c) can be calculated using (4.1), where A is the area of the binary image, n is the image width and m is the image height [29].

$$X_c = \frac{\sum_n x_n}{A} \quad \text{and} \quad Yc = \frac{\sum_m y_m}{A} \tag{4.1}$$

4.2.3 Perimeter

The perimeter is another geometrical parameter, which can easily be obtained from the chain code of the mass boundary. It is only needed to count the length of the chain code and take into consideration that steps in diagonal directions are longer by a factor of $\sqrt{2}$. The perimeter p is then given by an 8-neighborhood chain code is given in (4.2).

$$p = n_e + \sqrt{2} n_o \tag{4.2}$$

where n_e and n_o are the number of even and odd chain code steps, respectively. In contrast to the area, the perimeter is sensitive to noise level present in the image. The noisier the image is, the more rugged and thus longer the boundary of an object will become in the segmentation procedure.

4.2.4 Normalized Circularity

Area and perimeter are two parameters, which describe the size of a mass in one or the other way. In order to compare masses, which are observed from different distances, it is important to use shape parameters that do not depend on the size of the mass on the image plane. The circularity c is one of the simplest parameters of this kind. It is defined as:

$$c = \frac{p^2}{A} \tag{4.3}$$

The circularity is a dimensionless number with a minimum value of $4\pi \cong 12.57$ for circles. The circularity is 16 for a square and 20.8 for an equilateral triangle. Generally this parameter shows large values for elongated masses. In addition to this, normalized

circularity, defined in (4.4), equals zero for circles and tends to unity for complex shapes [30].

$$c_N = 1 - \frac{4\pi A}{p^2} \tag{4.4}$$

4.2.5 Radial Distance, Mean and Standard Deviation

The radial distance is measured by first detecting the centroid of the mass. The Euclidean distance from the centroid to the edge is then measured for the entire boundary. An arbitrary starting point is chosen and the boundary is followed clockwise. The radial distance is computed using (4.5).

$$d(i) = \sqrt{(x(i) - Xc)^2 + (y(i) - Yc)^2} \tag{4.5}$$

where (Xc, Yc) are the coordinates of the centroid, $(x(i), y(i))$ are the coordinates of the boundary pixel at the ith location.

The radial distance mean is computed using (4.6)

$$d_{avg} = \frac{1}{N}\sum_{i=1}^{N} d(i) \tag{4.6}$$

In the above equation N representss the number of boundary pixels of the extracted region. In the present study, N = 360. The standard deviation of the radial distance is computed using (4.7).

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (d(i) - d_{avg})^2} \tag{4.7}$$

4.2.6 Area Ratio

The area ratio parameter is defined in (4.8).

$$A_{ratio} = \frac{1}{d_{avg} * N}\sum_{i=1}^{N} (d(i) - d_{avg}) \tag{4.8}$$

4.3 Moment Based Shape Features

In probability theory, statistical moments are used to characterize statistical properties of random variables, e.g., expectations, variances, covariances and even probability density functions can be expressed in terms of statistical moments of various orders. Replacing the density function with a binary version, the same definitions can be used to obtain practical and useful set of shape descriptors. The moments of order $p+q$ of a region R represented by the bitmap $b_{n,m}$ can be calculated using (4.9).

$$M_{p,q} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} n^p m^q b_{n,m} = \sum_{n,m \in R} n^p m^q \tag{4.9}$$

The zero order moment of a density function is the volume covered by that function. In bitmap definition, the altitude of the function is either one or zero, implying that volume coincides with the area of the region. In fact $M_{0,0}$ is the number of pixels inside the region, which is also equal to the area of R in pixel units. The first order moments $M_{0,1}$ and $M_{1,0}$ are related to the balance point (x_c, y_c) of the region. The point (x_c, y_c) is also called the *center of gravity* or *centroid* and can be calculated using (4.10).

$$x_c = \frac{M_{1,0}}{M_{0,0}} \quad \text{and} \quad y_c = \frac{M_{0,1}}{M_{0,0}} \tag{4.10}$$

In order to make the description independent of position, moments can be calculated with respect to the centroid. These lead to the *central moments* and can be computed using (4.11).

$$\mu_{p,q} = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} (n-x_c)^p (m-y_c)^q b_{n,m} = \sum_{n,m \in region} (n-x_c)^p (m-y_c)^q \tag{4.11}$$

If the ordinary moments are known, it is less computationally expensive to derive the central moments from the ordinary moments than to evaluate (4.11) directly. For instance;

$$\mu_{0,0} = M_{0,0}$$

$$\mu_{0,1} = \mu_{0,1} = 0$$

$$\mu_{0,2} = M_{2,0} - x_c * M_{1,0}$$

$$\mu_{1,1} = M_{1,1} - x_c * M_{0,1}$$

4.3.1 Orientation

The second order central moments exhibit a number of properties that are comparable with covariance matrices in probability theory and the moments of inertia associated with rotating bodies in mechanisms. The principal axes of a region are spanned by the eigenvectors of the matrix, includes $\mu_{2,0}$, $\mu_{1,1}$, $\mu_{1,1}$, $\mu_{0,2}$. The principal moments are the corresponding eigenvalues calculated with (4.12.a) and (4.12.b).

$$\lambda_{max} = \frac{1}{2}(\mu_{2,0} + \mu_{0,2}) + \frac{1}{2}\sqrt{\mu_{2,0}^2 + \mu_{0,2}^2 - 2\mu_{0,2}\mu_{2,0} + 4\mu_{1,1}^2} \quad (4.12.a)$$

$$\lambda_{min} = \frac{1}{2}(\mu_{2,0} + \mu_{0,2}) + \frac{1}{2}\sqrt{\mu_{2,0}^2 + \mu_{0,2}^2 - 2\mu_{0,2}\mu_{2,0} + 4\mu_{1,1}^2} \quad (4.12.b)$$

The direction of the largest principal moment is measured with (4.13). θ is often used to specify the orientation of a region.

$$\theta = \tan^{-1}\left(\frac{\lambda_{max} - \mu_{2,0}}{\mu_{1,1}}\right) \quad (4.13)$$

4.3.2 Eccentricity

The eccentricity of a region can be defined as the ratio between square roots of the two principal moments and calculated with (4.14). The parameter depends solely on the shape, not on size and orientation.

$$eccentricity = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \quad (4.14)$$

4.3.3 Moment Invariants

Since $\mu_{0,0} = M_{0,0}$ is the area of the region, it can be used as a measure of size. As such it is useful in order to normalize the moments so as to arrive at a size independent description. The normalized central moments can be computed by (4.15).

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^{\frac{p+q}{2}+1}} \tag{4.15}$$

From the normalized moments up to order three, seven parameters can be extracted called moment invariants as shown below;

$$h_1 = \eta_{2,0} + \eta_{0,2}$$

$$h_2 = (\eta_{2,0} - \eta_{0,2})^2 + 4*\eta_{1,1}$$

$$h_3 = (\eta_{3,0} - 3*\eta_{1,2})^2 + (3*\eta_{2,1} - \eta_{0,3})^2$$

$$h_4 = (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2$$

$$h_5 = (\eta_{3,0} - 3*\eta_{2,1})*(\eta_{3,0} + \eta_{1,2})*[(\eta_{3,0} + \eta_{1,2})^2 - 3*(\eta_{2,1} + \eta_{0,3})^2]$$
$$+ (3*\eta_{2,1} - \eta_{0,3})*(\eta_{2,1} + \eta_{0,3})*[3*(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2]$$

$$h_6 = (\eta_{2,0} - \eta_{0,2})*[(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2]$$
$$+ 4*\eta_{1,1}*(\eta_{3,0} + \eta_{1,2})*(\eta_{2,1} + \eta_{0,3})$$

$$h_7 = (3*\eta_{2,1} - \eta_{0,3})*(\eta_{0,3} + \eta_{1,2})*[(\eta_{3,0} + \eta_{1,2})^2 - 3*(\eta_{2,1} + \eta_{0,3})^2]$$
$$+ (3*\eta_{2,1} - \eta_{0,3})*(\eta_{2,1} + \eta_{0,3})*[3*(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2]$$

This set of moments is invariant to translation, rotation and scale change [31].

4.4 Fourier Descriptors

The description of mass shape by moments is related to the area because it uses all the pixels within the masses. Since the shape of a mass is entirely described by its boundary, an alternative possibility for shape analysis is via the Fourier descriptors, since they use only the boundary.

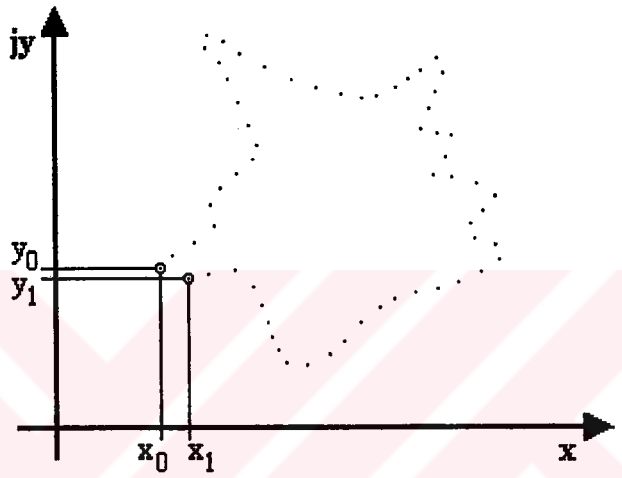

Figure 4.3 A digital boundary and its representation as a complex sequence.

The x-axis is treated as the real axis and the y-axis represents the imaginary axis of a sequence of complex numbers as shown in Figure 4.3. This representation has two advantages; it requires much less computational effort and reduces a 2-D to a 1-D problem [32]. The discrete Fourier transform of $s(i)$ is computed via (4.17) using (4.16) and (4.17).

$$s(i) = x(i) + jy(i) \tag{4.16}$$

$$a(u) = \frac{1}{N}\sum_{i=0}^{N-1} s(i)\exp(-j2\pi ui/N) \tag{4.17}$$

Fourier descriptors are not directly sensitive to rotation, scaling and starting point; these changes can be related to simple transformations on the descriptors as expressed in Table 4.1.

Table 4.1
Basic properties of Fourier descriptors.

Transformation	Boundary	Fourier Descriptor
Identity	$s(i)$	$a(u)$
Rotation	$s_r(i) = s(i)*e^{j\theta}$	$a_r(u) = a(u)*e^{j\theta}$
Scaling	$s_s(i) = \alpha s(i)$	$a_s(u) = \alpha a(u)$
Starting Point	$s_p(i) = s(i - i_0)$	$a_p(u) = a(u)*e^{-2ji_0^{u/N}}$

To avoid sensitivity problems, before the computation[1] of $a(u)$, mammographic mass shapes are normalized according to their rotation and scaling, using the orientation and area parameters computed as explained in the previous section. A fixed starting point is chosen for all shapes as shown in Figure 4.4.

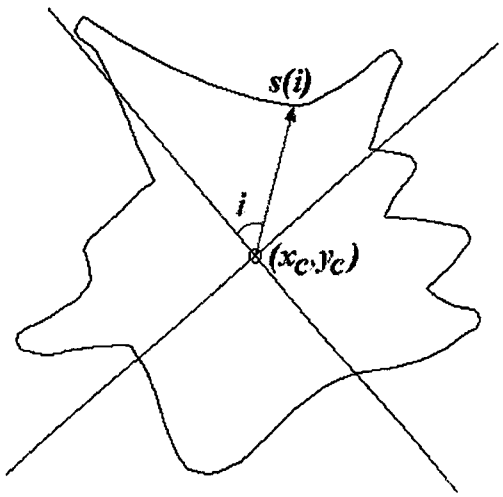

Figure 4.4 Normalization for Fourier descriptors.

[1] For 10 harmonics and 360 points (N = 360 denotes 1 degree increment).

5. CLASSIFYING MAMMOGRAPHIC MASS SHAPES

5.1 Introduction

The application of shape analysis has been applied to computerized mammographic methods. Magnin *et al.* and Davies and Dance, and Shen *et al.* utilized shape descriptors, which included various measures of area, compactness, eccentricity, and convexity [33]. However their methods were not applied to the classification of masses. The uses of shape features to classify breast masses has also been investigated. Luo *et al.* investigated further the use of shape descriptors which were simply difference measures between a tumor model and the measured mass. For this system, the shape features included area and average radial distance [34].

On the other hand, Brzakovic develop a system, which also used difference measures between the shape of tumor models and observed masses. In order to differentiate between round and stellate masses, a model of round benign tumors was developed. The least mean square difference was computed for the measured signature and model functions. The minimum difference was used as one of the classifier. Other features used in the classification process were area, edge distance variations, and edge intensity variations [35]. But it is important to point that these measures did not specifically characterize the mammographic mass shape.

Recently, Kilday et al. reported a comprehensive study of methods using shape descriptors as the main means of classification. Their study included the classification of breast masses into categories. Their optimum method used patient age and some shape descriptors in the classification process [36].

Several qualitative and quantitative techniques have been developed for characterizing the shape of masses in an image. These techniques are useful for classifying masses in a pattern recognition system and for symbolically describing masses in an image understanding system. Because of the malignant mass pathology, the shape of the mammographic mass can be used to discriminate between malignant and benign masses.

In this study geometric parameters such as area, perimeter, circularity, normalized circularity, radial distance mean and standard deviation, area ratio, orientation, eccentricity, moment invariants and Fourier descriptors up to 10, are calculated.

5.2 Feature Extraction

In the first step, an expert radiologist segments the mammographic mass shapes, in the training data set consisting of a total of 60 cases including 30 benign and 30 malignant. Geometric parameters of the pre-segmented mammographic mass shapes within this training set are then automatically computed using software specially developed for this purpose. A careful examination of the database and experimentation with various classification techniques show that, normalized circulatory, area and Fourier coefficients can be used successfully to classify masses as benign or malignant. This has been described in the following section.

5.3 Classification Scheme

Pattern recognition applications come in many forms. In some instances, there is an underlying and quantifiable statistical basis for the generation of patterns. In other instances, the underlying structure of the pattern provides the information fundamental for pattern recognition. In still others, neither of the above cases hold true, but it is possible to develop and "train" an architecture to correctly associate input patterns with desired responses. Thus, a given problem may allow one or more of these different solution approaches. When faced with solving a problem, one of the guiding principles an engineer or a scientist should apply is to, loosely speaking, "use the right tool for the job!".

Statistical pattern recognition assumes a statistical basis for classification of algorithms. A set of characteristic measurements, denoted features, is extracted from the input data and are used to assign each feature vector to one of c classes. Features are assumed generated by a state of nature, and therefore the underlying model is of a state of nature or class-conditioned set of probabilities and/or probability density functions.

Many times the significant information in a pattern is not merely in the presence or absence, or the numerical values, of a set of features. Rather, the interrelationships or interconnections of features yield important structural information, which facilitates structural description or classification. This is the basis of syntactic (or structural) pattern recognition. However, in using syntactic pattern recognition approaches, it is needed to be able to quantify and extract structural information and to assess structural similarity of patterns. One syntactic approach is to relate the structure of patterns with the syntax of a formally defined language, in order to capitalize on the vast body of knowledge related to pattern (sentence) generation and analysis (parsing). Typically, syntactic pattern recognition approaches formulate hierarchical descriptions of complex patterns built up from simpler sub-patterns. At the lowest level, primitive elements or "building blocks" are extracted from the input data. One discrepancy or distinguishing characteristic of syntactic patter recognition involves the choice of primitives. Primitives must be sub-patterns or building blocks, whereas features are any measurements

Modern digital computers do not emulate the computational paradigm of biological systems. The alternative of neural computing emerged from attempts to draw on knowledge of how biological neural systems store and manipulate information. This leads to a class of artificial neural systems termed neural networks. This study involves an amalgamation of research in many diverse fields such as psychology, neuroscience, cognitive science, and systems theory, and has recently received considerable renewed worldwide attention.

Neural networks are a relatively new computational paradigm. It is probably safe to say that the advantages, disadvantages, applications, and relationships to traditional computing are not fully understood. The notion that artificial neural networks can solve all problems in automated reasoning, or even all pattern recognition problems, is probably unrealistic. The structures of all approaches described above are shown in Figure 5.1.a, 5.1.b and 5.1.c.

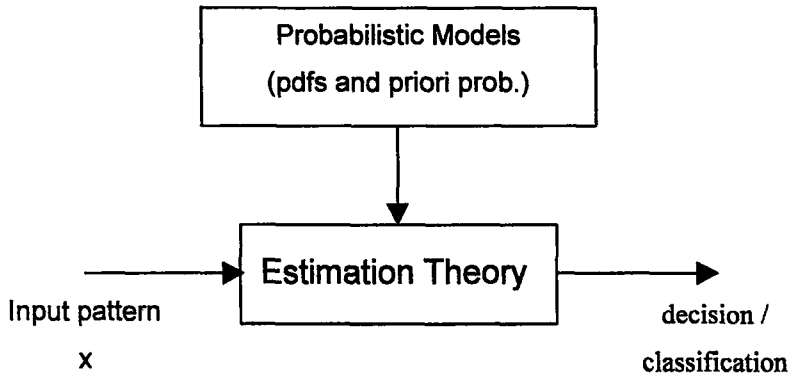

Figure 5.1.a Statistical pattern recognition approach.

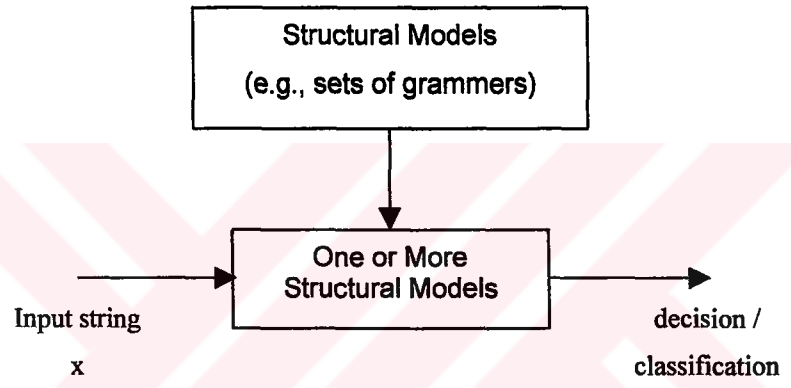

Figure 5.1.b Syntactic pattern recognition approach.

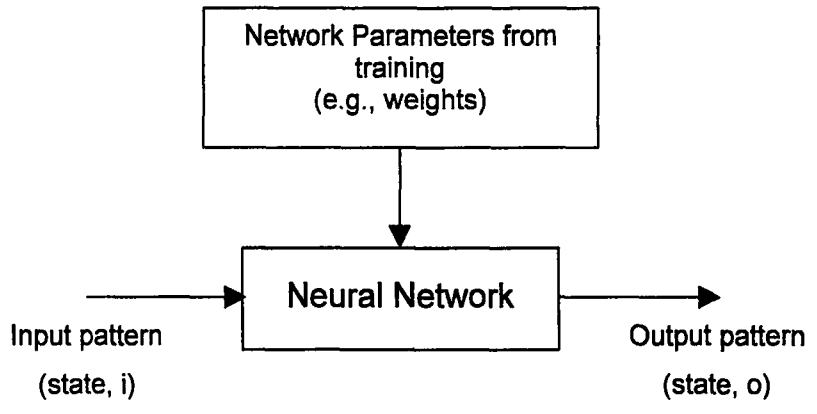

Figure 5.1.c Neural network approach.

The boundaries between statistical, syntactic and neural pattern recognition are fuzzy and fading. They share common features and common goals. Often, given a specific pattern recognition problem, one approach over another may be chosen based on an analysis of underlying *statistical* components (statistical pattern recognition), underlying *grammatical* structure (syntactic pattern recognition), as well as suitability of a *neural network* solution (and training ability), and perhaps lack of suitable statistical or structural models.

The structure of the pattern is often deemed insignificant in the statistical pattern recognition or decision theoretic approach. However, the structure could be reflected by a suitable choice of features (e.g., a binary feature vector could indicate the presence or absence of observed relations). Similarly, the neural pattern recognition approach is, in some cases, an implementation derived from statistical and syntactic pattern recognition approaches. When explicit structural information about the patterns is available, it makes sense to choose syntactic pattern recognition. When this information is either unavailable or irrelevant, then statistical pattern recognition may be used, that is why, a statistical method is used for this study.

In statistical pattern recognition, developing decision or classification strategies, which form classifiers, are concerned. Classifier design attempts to integrate all available problem information, such as measurements and a priori probabilities. Decision rules may be formulated in several interrelated ways, for example:

- By converting an a priori class probability $P(w_i)$ into a measurement-conditioned ("a posteriori") probability $P(w_i \mid \underline{x})$.

- By formulating a measure of expected classification error or "risk", and choosing a decision rule that minimizes this measure.

Both strategies lead to a partitioning of $\mathbf{R}^d$ and may be implemented via discriminant functions. For a single measurement statistical model, a case is considered where dimension of feature space $d = 1$, number of classes $c = 2$, $P(w_1) = P(w_2)$ and

$$p(x \mid w_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma} \right)^2 \right] \tag{5.1}$$

This is already a somewhat specialized case, since the variance and prior probabilities for both classes are the same. Thus, only the class means provide class-specific information assuming that $\mu_1 = \mu_2$. It is observed that Bayes rule yields

$$P(w_i \mid x) = p(x \mid w_i)\left[\frac{P(w_i)}{p(x)} \right] \tag{5.2}$$

where $p(x)$ is the unconditional density function. Therefore the decision rule is adopted as "choose w_1 if $p(x \mid w_1) > p(x \mid w_2)$, otherwise choose w_2." This leads to the decision regions defined by α, as shown in Figure 5.2.

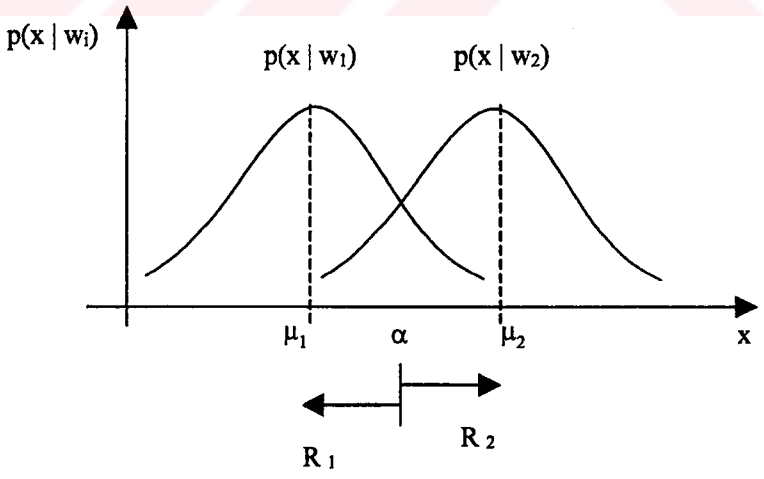

Figure 5.2 Densities for a single measurement.

In developing models that characterize the class-specific generation of the feature vectors, the correct role of a density function must be carefully observed. Density functions are characterizations of random vectors; they do not generate random vectors. To enable a rigorous solution of the problem, Bayes theorem is used.

5.3.1 Bayes Theorem and Bayes Risk Criterion

Bayes theorem (Figure 5.3), enables computation of the a priori estimate of the probability of a certain class in terms of the a posteriori or measurement conditioned, probability of a state of nature via (5.3) and (5.4), as given in (5.3).

$$P(w_i \mid \underline{x}) = \frac{[p(\underline{x} \mid w_i)P(w_i)]}{p(\underline{x})} \tag{5.3}$$

where

$$p(\underline{x}) = \sum_i p(\underline{x} \mid w_i)P(w_i) \tag{5.4}$$

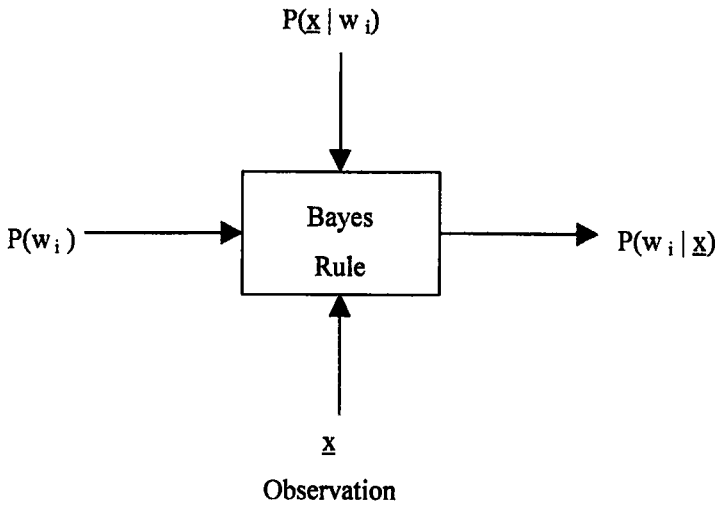

Figure 5.3 Converting $P(w_i)$ to $P(w_i \mid \underline{x})$ via $\underline{x}$.

The Bayes decision criterion employs a systematic procedure of assigning a cost to each correct and incorrect decision and then minimizing the average cost. Let C_{jk} be the cost of making decision d_j when m_k is true, then for the binary decision problem there are four possible costs [37]:

C_{11}, cost of deciding d_1 when m_1 is true.

C_{12}, cost of deciding d_1 when m_2 is true.

C_{21}, cost of deciding d_2 when m_1 is true.

C_{22}, cost of deciding d_2 when m_2 is true.

If a natural paring of d_j and m_k is assumed then it is noted that C_{11} and C_{22} are costs associated with correct decisions while C_{12} and C_{21} are associated with incorrect decisions.

It may seem strange to talk of a cost associated with a correct decision. Although C_{11} and C_{22} are often set to zero, there is nothing inconsistent in assigning a cost to a correct decision. For example, the cost of making a correct medical diagnosis might be the risks involved in the treatment. The expected or average cost is given in (5.5).

$$\begin{aligned} B = E\{C_{jk}\} = C_{11} * P\{d_1, m_1\} + C_{12} * P\{d_1, m_2\} \\ + C_{21} * P\{d_2, m_1\} + C_{22} * P\{d_2, m_2\} \end{aligned} \tag{5.3}$$

Since $P\{d_j, m_k\} = P\{d_j \mid m_k\} * P\{m_k\}$, (5.3) can take the form as in (5.4)

$$\begin{aligned} B = (C_{11} * P\{d_1 \mid m_1\} + C_{12} * P\{d_1 \mid m_2\}) * P\{m_1\} \\ + (C_{21} * P\{d_2 \mid m_1\} + C_{22} * P\{d_2 \mid m_2\}) * P\{m_2\} \end{aligned} \tag{5.4}$$

The average cost also be written as in (5.5).

$$\begin{aligned} B = E\{C_{jk}\} = E\{C_{jk} \mid m_1\} * P\{m_1\} + E\{C_{jk} \mid m_2\} * P\{m_2\} \\ = B_1 * P\{m_1\} + B_2 * P\{m_2\} \end{aligned} \tag{5.5}$$

B_1 and B_2 are referred to as conditional costs; B_k is the average cost assuming that m_k is true and can be quickly computed with (5.6.a) and (5.6.b).

$$P\{d_1 \mid m_1\} = 1 - P\{d_2 \mid m_1\} \tag{5.6.a}$$

$$P\{d_1 \mid m_2\} = 1 - P\{d_2 \mid m_2\} \tag{5.6.b}$$

Then the conditional costs can be written as in (5.7.a) and (5.7.b).

$$B_1 = C_{11} + (C_{21} - C_{11}) * P\{d_2 \mid m_1\} \tag{5.7.a}$$

$$B_2 = C_{12} + (C_{12} - C_{22}) * P\{d_2 \mid m_2\} \tag{5.7.b}$$

Therefore the average cost is given by (5.8) and (5.9).

$$\begin{aligned} B = {} & C_{11} * P\{m_1\} + (C_{21} - C_{11}) * P\{d_2 \mid m_1\} * P\{m_1\} \\ & + C_{12} * P\{m_2\} - (C_{12} - C_{22}) * P\{d_2 \mid m_2\} * P\{m_2 \end{aligned} \tag{5.8}$$

$$\begin{aligned} B = {} & C_{11} * P\{m_1\} + C_{12} * P\{m_2\} \\ & + (C_{21} - C_{11}) * P\{m_1\} \int_{Z2} p(z \mid m_1) dz - (C_{12} - C_{22}) * P\{m_2\} \int_{Z2} p(z \mid m_2) dz \end{aligned} \tag{5.9}$$

Combining the two integrals gives (5.12).

$$\begin{aligned} B = {} & C_{11} * P\{m_1\} + C_{12} * P\{m_2\} \\ & + \int_{Z_2} [(C_{21} - C_{11}) * P\{m_1\} p(z \mid m_1) - (C_{12} - C_{22}) * P\{m_2\} p(z \mid m_2)] dz \end{aligned} \tag{5.10}$$

Bayes decision criterion states to select the decision region Z_2 in order to minimize the average cost B.

In this study, C_{11} and C_{22} costs are set to zero while C_{12} and C_{21} costs are selected specifically to hit more malignant cases. Then the decision regions, where average costs are minimized, are obtained. A decision tree, consisting of the Bayesian classifiers as shown in Figure 5.4, is implemented to classify both of the two cases as errorless as possible. During the selection of mammographic mass shape parameters in Bayesian classifiers ROC analysis is used.

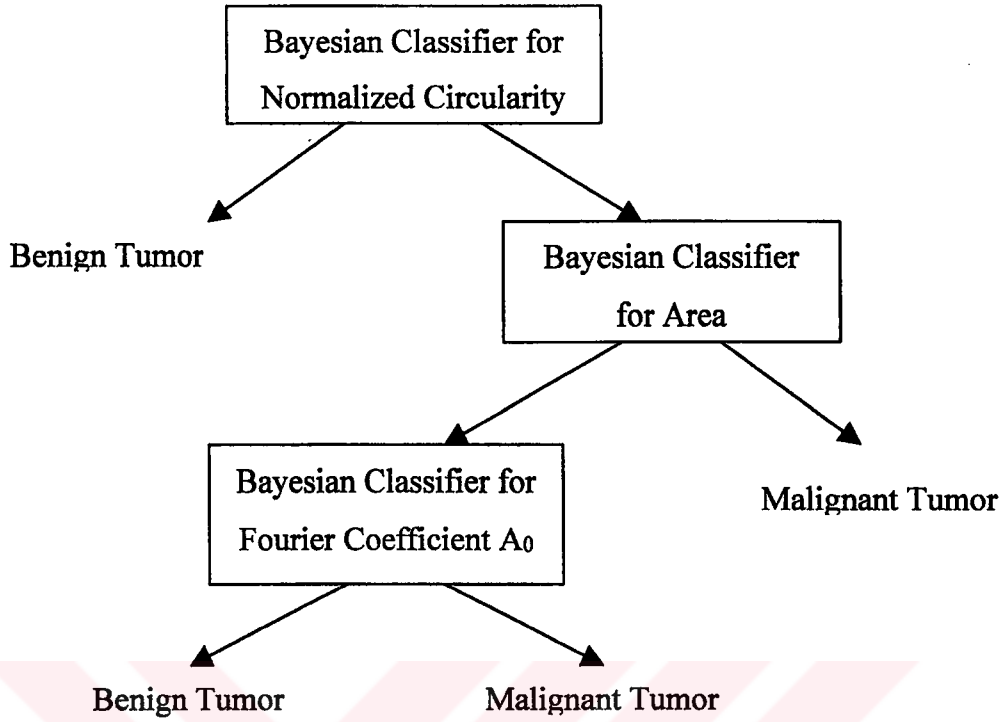

Figure 5.4 The decision tree for classification.

5.3.2 ROC Analysis

ROC analysis is part of a field called "Signal Detection Theory" developed during World War II for the analysis of radar images [38]. ROC analysis is now common in medicine, healthcare and particularly in radiology, where it is used to quantify the accuracy of diagnostic tests [39].

ROC analysis is the standard approach to evaluate the sensitivity and specificity of diagnostic procedures. ROC analysis estimates a curve, which describes the inherent tradeoff between sensitivity and specificity of a diagnostic test. Each point on the ROC curve is associated with a specific diagnostic criterion. This point will vary among observers because their diagnostic criteria will vary even when their ROC curves are the same [40]. A ROC curve shows the tradeoff between sensitivity and specificity. Any increase in sensitivity will be accompanied by a decrease in specificity. The closer the

curve follows the left-hand border and then the top border of the ROC space, the more accurate the test. The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test [41].

The area under the ROC curve (A-z) has become a particularly important metric for evaluating diagnostic procedures because it is the average sensitivity over all possible specificities [42, 43]. Accuracy is measured by the area under the ROC curve. An area of 1 represents a perfect test; an area of 0.5 represents a worthless test [44]. Three ROC curves representing excellent, good, and worthless tests are shown in Figure 5.5.

A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system given in Table 5.1.

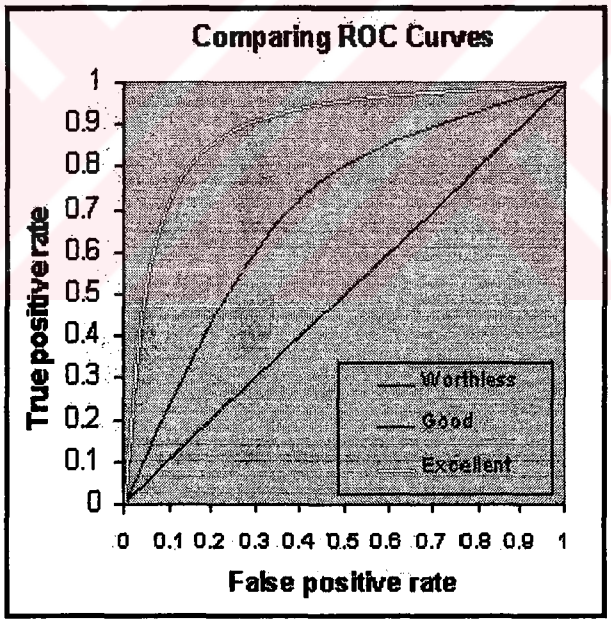

Figure 5.5 Different ROC curves.

Table 5.1
Traditional academic point system for accuracy.

Area under the curve	Accuracy
> .90	excellent (A)
> .80 and < .90	good (B)
> .70 and < .80	fair (C)
> .60 and < .70	poor (D)
> .50 and < .60	fail (F)

Two methods are commonly used to compute the area under the ROC curve [45]:

- *A non-parametric method* based on constructing trapezoids under the curve as an approximation of area.
- *A parametric method* using a maximum likelihood estimator to fit a smooth curve to the data points.

ROC analysis and the area under the curve for each mammographic shape parameter computed by the software show that normalized circularity parameter is the best for the first decision, while area is the best for the second decision and the Fourier coefficient A_0 is the best for the last decision in the designed decision tree structure.

5.3.3 Classification Scheme Performance

The developed computer software has been redesigned to implement the decision tree. A total number of 25 biopsy-performed mammographic mass shapes have been analyzed by the software and decisions have been made listed in Table 5.2.

Table 5.2
Decisions of the software.

		Biopsy Result	
		Malignant	Benign
Decision	Malignant	9	1
	Benign	1	14

6. MAMMOGRAM DATABASE DEVELOPMENT

6.1 Introduction

Database systems store information in every conceivable health care environment. From large tracking databases such as hospital information systems to a patient's examination report, database systems store and distribute the data that are depended on. Until the last few years, large database systems could be run only on large mainframe computers. These machines have traditionally been expensive to design, purchase, and maintain. However, today's generation of powerful, inexpensive workstation computers enables programmers to design software that maintains and distributes data quickly and inexpensively.

6.2 Relational Database Management System

The most popular data storage model is the relational database, which grew from the seminal paper "A Relational Model of Data for Large Shared Data Banks," written by Dr. E. F. Codd in 1970. Dr. Codd defined 13 rules, oddly enough referred to as Codd' s 12 Rules, for the relational model:

- A relational DBMS must be able to manage databases entirely through its relational capabilities.
- All information in a relational database, including table and column names, is represented explicitly as values in tables.
- Every value in a relational database is guaranteed to be accessible by using a combination of the table name, primary key value, and column name.
- The DBMS provides systematic support for the treatment of null values (unknown or inapplicable data), distinct from default values and independent of any domain.
- The description of the database and its contents is represented at the logical level as tables and can therefore be queried using the database language.

- At least one supported language must have a well-defined syntax and be comprehensive. It must support data definition, manipulation, integrity rules, authorization and transactions.
- All views that are theoretically updateable can be updated through the system.
- The DBMS supports not only set-level retrievals but also set-level inserts, updates and deletes.
- Application programs are logically unaffected when physical access methods or storage structures are altered.
- Application programs and ad hoc programs are logically unaffected, to the extent possible, when changes are made to the table structures.
- The database language must be capable of defining integrity rules. They must be stored in the online catalog, and they cannot be bypassed.
- Application programs and ad hoc requests are logically unaffected when data is first distributed or when it is redistributed.
- It must not be possible to bypass the integrity rules defined through the database language by using lower-level languages.

Most databases have had a "parent/child" relationship; that is, a parent node would contain file pointers to its children as seen in Figure 6.1.

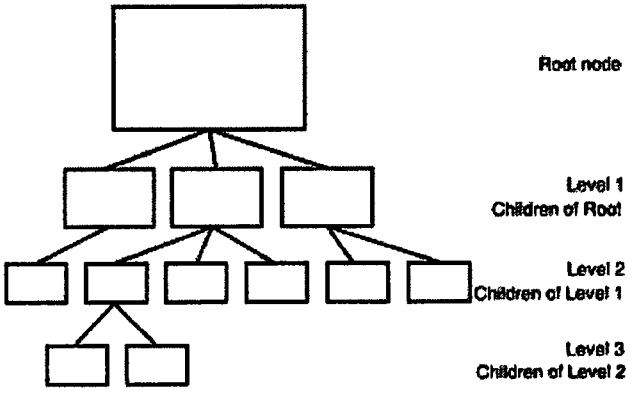

Figure 6.1 Codd' s relational database management system.

This method has several advantages and many disadvantages. In its favor is the fact that the physical structure of data on a disk becomes unimportant. The programmer simply stores pointers to the next location, so data can be accessed in this manner. Also, data can be added and deleted easily. However, different groups of information could not be easily joined to form new information. The format of the data on the disk could not be arbitrarily changed after the database was created. Doing so would require the creation of a new database structure.

Codd' s idea for an RDBMS uses the mathematical concepts of relational algebra to break down data into sets and related common subsets. Because information can naturally be grouped into distinct sets, Dr. Codd organized his database system around this concept. Under the relational model, data is separated into sets that resemble a table structure. This table structure consists of individual data elements called columns or fields. A single set of a group of fields is known as a record or row [46]. Structured query language evolved to service the concepts of the relational database model.

6.3 Structured Query Language (SQL)

The history of SQL begins in an IBM laboratory in San Jose, California, where SQL was developed in the late 1970s. The initials stand for Structured Query Language and the language itself is often referred to as "sequel." It was originally developed for IBM's DB2 product. SQL is a nonprocedural language, in contrast to the procedural or third-generation languages such as COBOL and C that had been created up to that time. Nonprocedural means what rather than how. SQL describes what data to retrieve, delete, or insert, rather than how to perform the operation and this property allows SQL to find an application area in the developed database.

SQL is the de facto standard language used to manipulate and retrieve data from relational databases. SQL enables a programmer or database administrator to do the followings:

- Modify a database's structure
- Change system security settings
- Add user permissions on databases or tables

- Query a database for information
- Update the contents of a database

The Select Statement

The most commonly used statement in SQL is the SELECT statement, which retrieves data from the database and returns the data to the user. In addition to the SELECT statement, SQL provides statements for creating new databases, tables, fields, and indexes, as well as statements for inserting and deleting records. ANSI SQL also recommends a core group of data manipulation functions.

The Insert...Values Statement

This statement enters data into a table one record at a time. It is useful for small operations that deal with just a few records. The syntax of this statement is as follows:

INSERT INTO table_name (col1, col2...) VALUES(value1, value2...)

The basic format of the insert statement adds a record to a table using the columns you give it and the corresponding values you instruct it to add. When inserting data into a table with this statement, three rules must be followed:

The values used must be the same data type as the fields they are being added to.

The data's size must be within the column's size. For instance, you cannot add an 80-character string to a 40-character column.

The data's location in the VALUES list must correspond to the location in the column list of the column it is being added to.

The UPDATE Statement

The purpose of the UPDATE statement is to change the values of existing records. The syntax is

UPDATE table_name SET columnname1 = value1 [, columname2 = value2] WHERE search_condition

This statement checks the WHERE clause first. For all records in the given table in which the WHERE clause evaluates to TRUE, the corresponding value is updated.

The DELETE Statement

In addition to adding data to a database, you will also need to delete data from a database. The syntax for the DELETE statement is

DELETE FROM tablename WHERE condition

6.4 Open Database Connectivity (ODBC)

ODBC is a functional library designed to provide a common Application Programming Interface (API) to underlying database systems. It communicates with the database through a library driver, just as Windows communicates with a printer via a printer driver. Depending on the database being used, a networking driver may be required to connect to a remote database. The architecture of ODBC is illustrated in Figure 6.2.

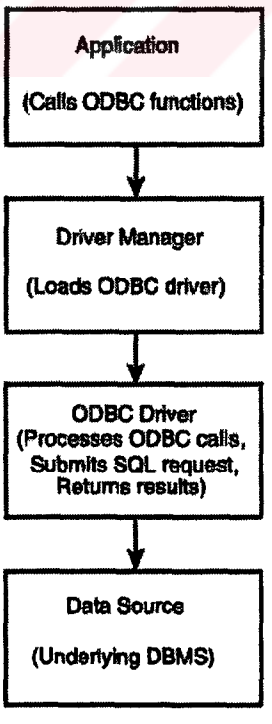

Figure 6.2 ODBC structure.

The unique feature of ODBC (as compared to the Oracle or Sybase libraries) is that none of its functions are database-vendor specific. ODBC has developed into a standard adopted into many products, including Visual Basic, Visual C++ and Borland Delphi.

6.5 Designing Mammogram Database

Microsoft Access is a PC-only database product that contains many of the features of a relational database management system and this is the one reason why it has been chosen for the mammogram database. The second reason for choosing Microsoft Access is the ODBC compatibility which enables internet access for telemedical use.

The most important decision for a database designer, after the hardware platform and the RDBMS have been chosen, is the structure of the tables. Decisions made at this stage of the design can affect performance and programming later during the development process. The developed mammogram database consists of mainly 5, total 9 numbers of tables. The tables named as *patientfile, reports, findings, biopsies* and *doctors* are mainly used to store electronic patient record. In addition to this, the tables labeled as *assessments, recommendations, reasons* and *biopsytech* are designed for the coding system of the software. During the design of these tables, much attention is focused on the American College of Radiology's Breast Imaging Reporting and Data System shortened as BI-RADS since it contains a guide to standard coding. The tables' relationships are as shown in Figure 6. 3.

6.6 Patient File Table

This table stores the morphologic data of the patient including patient's ID, photo and patient medical history related to breast cancer such as menopause and menarche ages, births before and after age 30, the prior breast cancers, ovarian and endothelial cancers are seen or not, patient' s mother and sister have the symptoms of breast cancer or not. The fields of the patient file table are as follows:

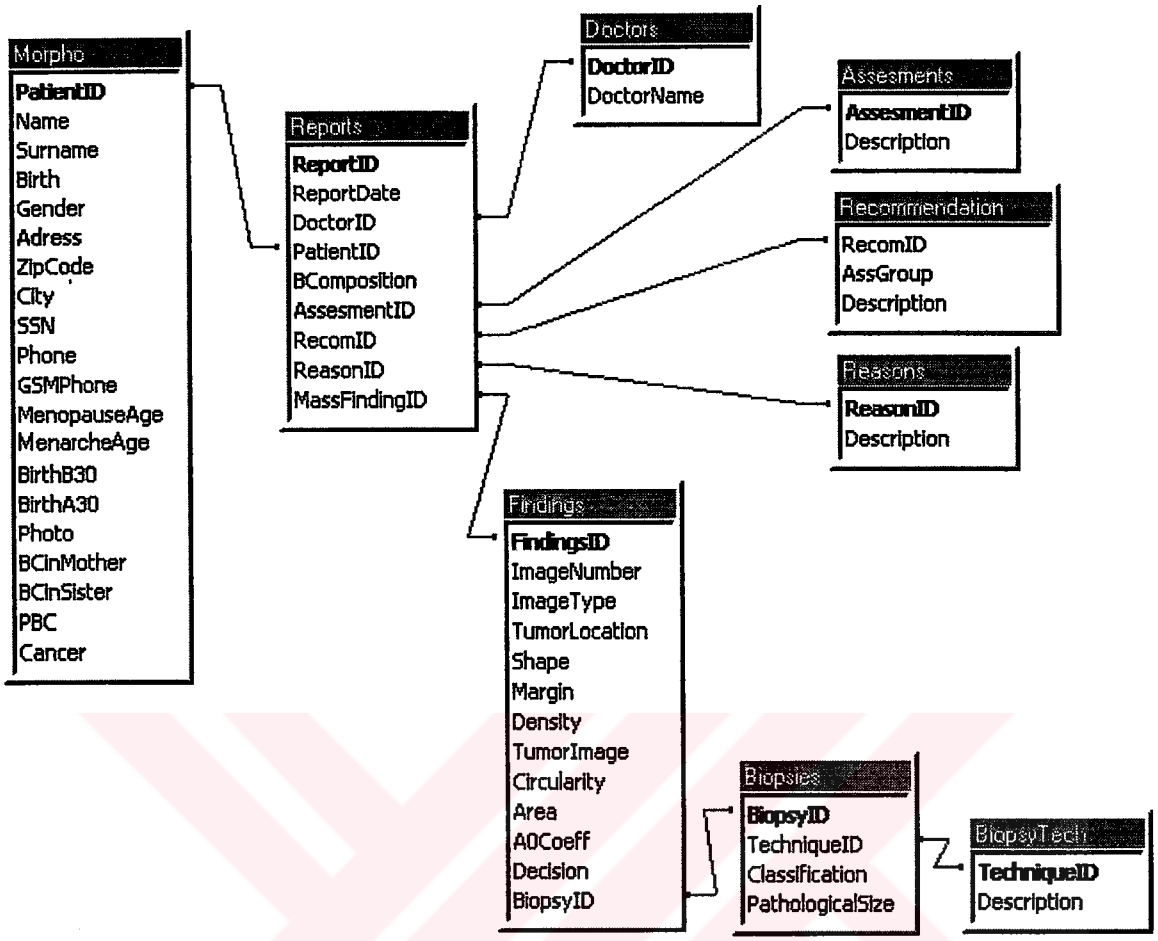

Figure 6.3 Relationships of tables in mammogram database.

- *Patient ID*, a 15 byte long string field, stores the TOM unique number of the patient.
- *Name*, a 20 byte long string field, stores the name of the patient.
- *Surname*, a 20 byte long string field, stores the surname of the patient.
- *Birth*, a date field, stores the birth date of the patient.
- *Gender*, a one byte long string field, stores the gender of the patient.
- *Address*, a 50 byte long string field, stores the postal address of the patient.
- *Zip Code*, a 5 byte long string field, stores the zip code for the address.
- *City*, a 2 byte long string field, stores the code of the city for the address.
- *SSN*, a 10 byte long string field, stores the social security number of the patient.
- *Phone*, a 11 byte long string field, stores the patient' s phone number.
- *GSM Phone*, a 11 byte long string field, stores the patient' s cellular phone number.

- *Photo*, A OLE object field, stores the patient photo.
- *MenopauseAge*, a 2 byte long string field, stores the menopause age of the patient.
- *MenarcheAge*, a 2 byte long string field, stores the menarchee age of the patient.
- *BirthB30*, a 2 byte long string field, stores the births before the age of 30.
- *BirthA30*, a 2 byte long string field, stores the births after the age of 30.
- *BCinMother*, a checkbox field, stores the information if breast cancer is detected in the patient' s mother or not.
- *BcinSister*, a checkbox field, stores the information if breast cancer is detected in the patient' s sister or not.
- *PBC*, a checkbox field, stores the information if breast cancer was detected previously in the patient or not.
- *Cancer*, a checkbox field, stores the information if other types of cancers were detected in the patient or not.

6.7 Reports Table

The examination date, the name of the doctor who examined, breast composition and the assessment, recommendation, examination reason codes are stored in this table. The fields of the reports table are as follows:

- *ReportID*, an auto incremental field, stores the report process number.
- *ReportDate*, a date field, stores the report date.
- *DoctorID*, a 2 byte long string field, stores the doctor code and permits relationship with doctors table.
- *PatientID*, a 15 byte long string field, stores the ID of patient and permits relationship with patientfile table.
- *Bcomposition*, a 1 byte long string field, stores 1 for entirely fat, 2 for fibroglandular densities, 3 for heterogeneously dense and 4 for extremely dense breast compositions.
- *AssesmentID*, a 1 byte long string field, stores the assessment code and permits relationship with assessments table.
- *RecomID*, a 1 byte long string field, stores the recommendation code and permits relationship with recommendations table.

- *ReasonID*, a 1 byte long string field, stores the code of the reason and permits relationship with reasons table.
- *MassFindingsID*, an auto incremental field, stores the biopsy process number and permits relationship with reports table.

6.8 Findings Table

Mammographic mass finding data, including the image and the location of the mass; normalized circularity, area and the Fourier coefficient A_0 parameters of the mass, software decision and the ID of the biopsy, if performed, are the fields of the findings table. The structure of this table is as follows:

- *FindingsID*, an auto incremental field, stores the biopsy process number and permits relationship with reports table..
- *ImageNumber*, a 15 byte long string, stores the mammographic film number where the mass is detected.
- *ImageType*, a 1 byte long string field, stores the view of projection.
- *Tumorlocation*, a 1 byte long string field, stores the location of the mass.
- *Shape*, a 1 byte long string field, stores 1 for round, 2 for oval, 3 for lobular and 4 for irregular mass shapes.
- *Margin*, a 1 byte long string field, stores 1 for circumscribed, 2 for microlobulated, 3 for obscured, 4 for indistinct and 5 for spiculated mass margins.
- *Density*, a 1 byte long string field, stores 1 for high, 2 for equal or isodense, 3 for low and 4 for fat containing density for the mass.
- *TumorImage*, an OLE field, stores the image of mammographic mass.
- *Circularity*, a number field, stored the computed normalized circularity.
- *Area*, a long number field, stored the computed area.
- *A0Coeff*, a number field, stored the computed Fourier coefficient A_0.
- *Decision*, a 1 byte long string code, stores B for benign and M for malignant decision done by the system.
- *BiopsyID*, an auto incremental field, stores the biopsy process number and permits relationship with biopsies table.

6.9 Biopsies Table

The biopsy data, consisting of the used technique, classification and the pathological size of the mass, are stored in this table. The fields of the biopsies table are as follows:

- *BiopsyID*, an auto incremental field, stores the biopsy process number and permits relationship with findings table.
- *TechniqueID*, a 2 byte long string field, stores the biopsy technique code.
- *Classification*, a 1 byte long string code, stores B for benign and M for malignant findings.
- *PathologicalSize*, a 10 byte long string field, stores the pathological tumor size.

6.10 Doctors Table

While the name of the doctor including surname are stored in this table, relationship with the reports table is obtained via DoctorID. The fields of this table are as follows:

- *DoctorID*, a 2 byte long string field, stores the code of a doctor.
- *DoctorName*, a 25 byte long string field, stores the name and the surname of the doctor.

6.11 Assessments Table

The assessment codes and their descriptions, listed in Table 6.1, defined in ACR' s BI-RADS, are stored in this table and the relationship with the reports table is obtained via AssesmentID. The fields of this table are as follows:

- *AssesmentID*, a 1 byte long string field, stores the codes for the assessments.
- *Description*, a 25 byte long string field, stores the description of each assessment code.

Table 6.1
Codes and descriptions for assessments as described in BI-RADS.

Assessment ID	Description
0	Needs additional imaging evaluation
1	Negative
2	Benign Finding
3	Probably Benign
4	Suspicious abnormality
5	Highly suggestive of malignancy

6.12 Recommendations Table

Proper recommendation codes for each assessment and the descriptions of the codes, as defined in ACR' s BI-RADS, are stored in the recommendations table. The constant data in this table are listed in Table 6.2 and the fields are as follows:

- *RecomID*, a 1 byte long string field, stores the proper recommendation codes for assessments.
- *AssGroup*, a 1 byte long string field, stores the assessment codes.
- *Description*, a 60 byte long string field, stores the description for the recommendation code.

Table 6.2
Codes for recommendations as defined in BI-RADS.

Recommendation ID	Assessment Group	Description
M	0	Magnification Views
S	0	Spot compression
V	0	Spot magnification views
U	0	Ultrasound
O	0	Old films for comparison
G	0	Ductography
P	0	Additional projections
D	1	Any decision to biopsy should be based on clinical assessment
N	1	Normal interval follow-up
D	2	Any decision to biopsy should be based on clinical assessment
N	2	Normal interval follow-up
F	3	Follow-up at short interval
B	4	Biopsy should be considered
L	4	Needle localization and biopsy
H	4	Histology using core biopsy
T	4	Suggestive of malignancy
Y	4	Cyctologic Analysis
Y	5	Cyctologic analysis
B	5	Biopsy should be strongly considered
L	5	Needle localization and biopsy
H	5	Histology using core biopsy
T	5	Highly suggestive of malignancy

6.13 Reasons Table

The reasons table consists of constant data, that are the reason codes and their descriptions as defined in ACR' s BI-RADS listed in Table 6.3.

Table 6.3
Reason codes and description for each code as defined in BI-RADS.

Reason ID	Description
A	Additional evaluation requested from outside
C	Clinical Finding
F	Follow-up at short interval from prior study
H	History of breast augmentation, asymtomatic
O	Review of outside study
R	Pre-reduction mammoplasty
S	Screening
T	Pre-radiation therapy
V	Additional evaluation requested

The ReasonID field permits the relationship with reports table and included fields are as follows:

- *ReasonID*, a 1 byte long string field, stores the reason code.
- *Description*, a 45 byte long string field, stores the reason code description.

6.14 BiopsyTech Table

This table includes the data of biopsy if performed. The codes of biopsy techniques, listed in Table 6.4, are as defined in ACR' s BI-RADS.. The fields of this table are as follows;

- *TecniqueID*, a 2 byte long string field, stores the code of the biopsy technique.
- *Description*, a 35 byte long string field, stores the description of the biopsy code.

Table 6.4
Biopsy technique codes and descriptions defined in BI-RADS.

Technique ID	Description
CA	Cyst Aspiration
CB	Core Biopsy
EB	Excisional Biopsy
MA	Mammog. Non-Stereotaxic Cyst Asp.
MR	MRI Guided Biopsy
SA	Streotactic Guided Cyst Aspiration
SB	Streotaxic Core Biopsy
UA	Ultrasound Guided Cyst Aspiration
UB	Ultrasound Guided Biopsy

6.15 Forms

After constructing the database structure, the next step is design of the forms. During this step, attention must be focused on the proper layout of data on the screen and helpful explanations about the results of user requests.

The first form appear on the computer screen is used to allow pre-defined users to access mammogram database. Purposing the access, username and password are asked The view of this form is shown in Figure 6.4

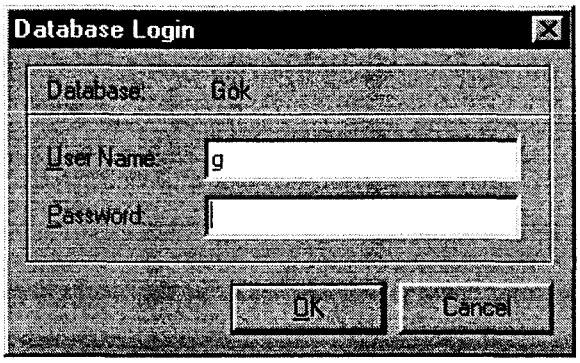

Figure 6.4 Database login form.

By entering the username and password correctly, the entrance form will be displayed on the screen and patient ID is asked. An automatic patient ID entering can be achieved by touching the patient TOM to the TOM reader connected to the computer via RS232. This event forces the software to start a patient data search process. At the end of the search if the patient record does not exit then the doctor is informed whether to create a new patient record or not as shown in Figure 6.5.

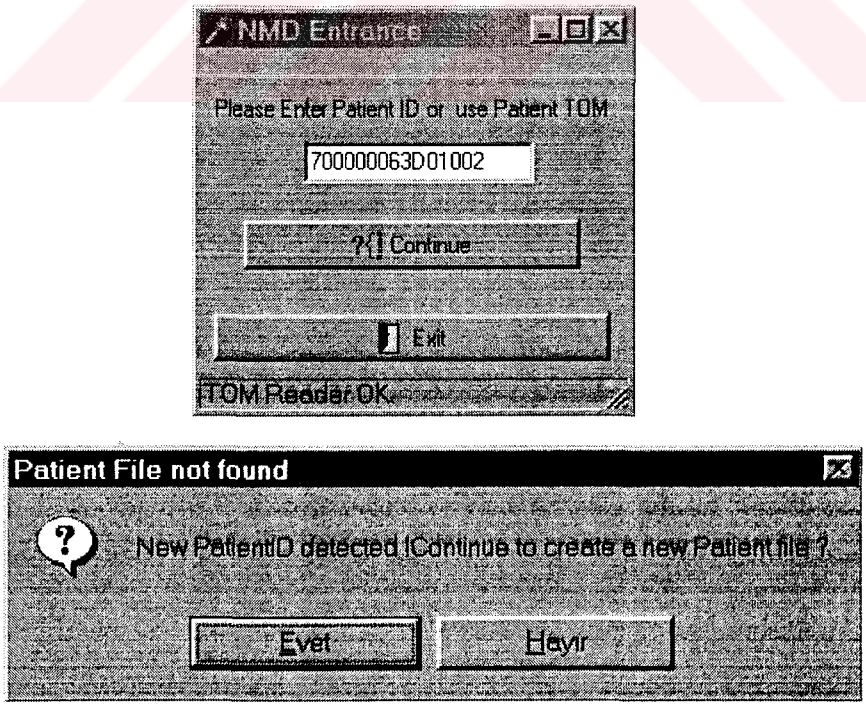

Figure 6.5 The case in which patient record could not found.

If the doctor accepts to create a new patient record then a new form appears on the screen to enter the morphological patient data including patient photo as shown in Figure 6.6.

Clicking the create new patient file button lead the system software to creates the new patient record and user is informed about the success of the request..

If the patient ID, read form the TOM via TOM reader, is found in the mammography database then a new form will be appeared on the screen. This form is the patient file form as shown in Figure 6.7 and it includes the morphological patient data and old mammographic reports list on the left side while the details of the report or new mammographic mass entry form on the right side.

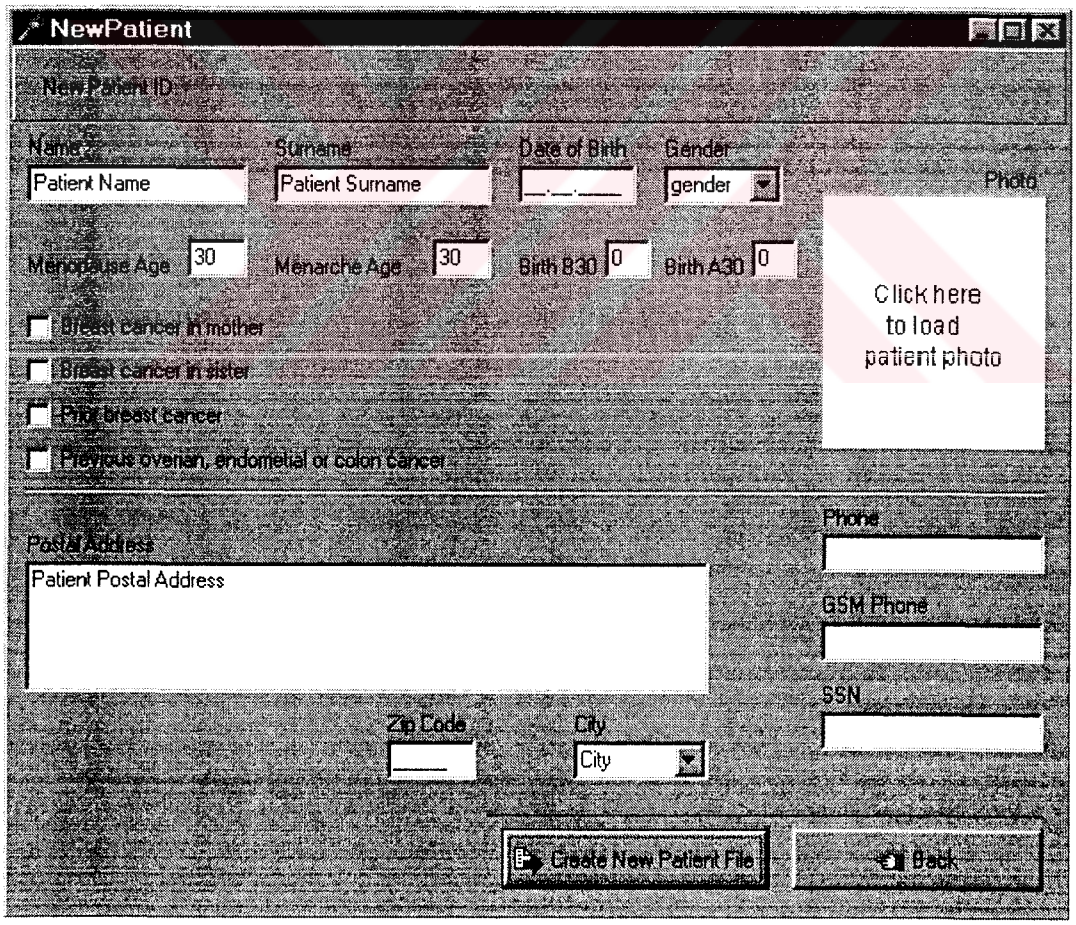

Figure 6.6 New patient data entry form.

Clicking the panel where click here to load mammographic mass is written down results of opening a picture dialog box to select the image of the mammographic mass. Selecting the image will force the software to start a computation process. The shape descriptors are computed to make a decision which is displayed in the decision panel as seen in Figure 6.8. If a biopsy is formed then it can also be stored

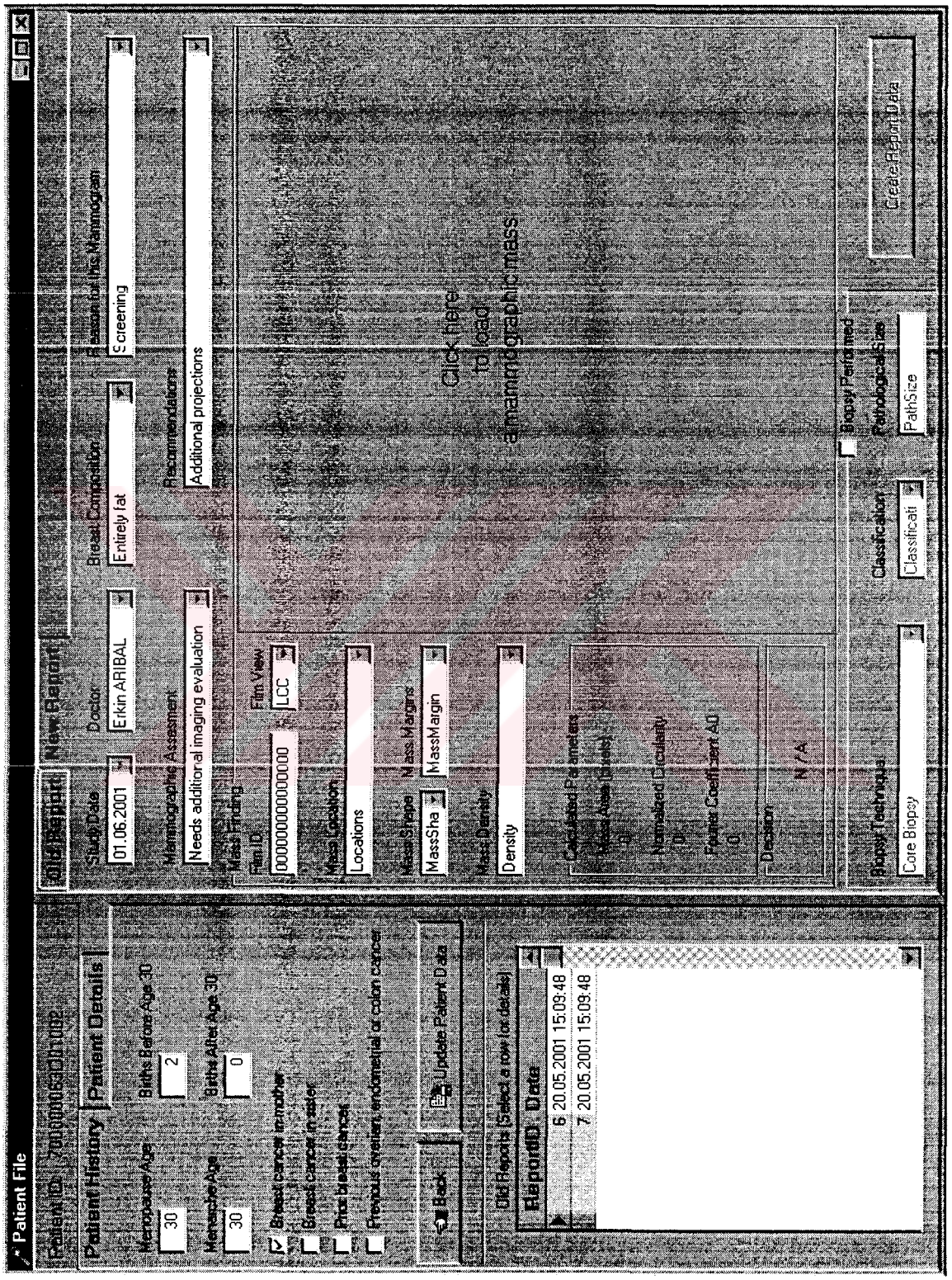

Figure 6.7 Patient file form.

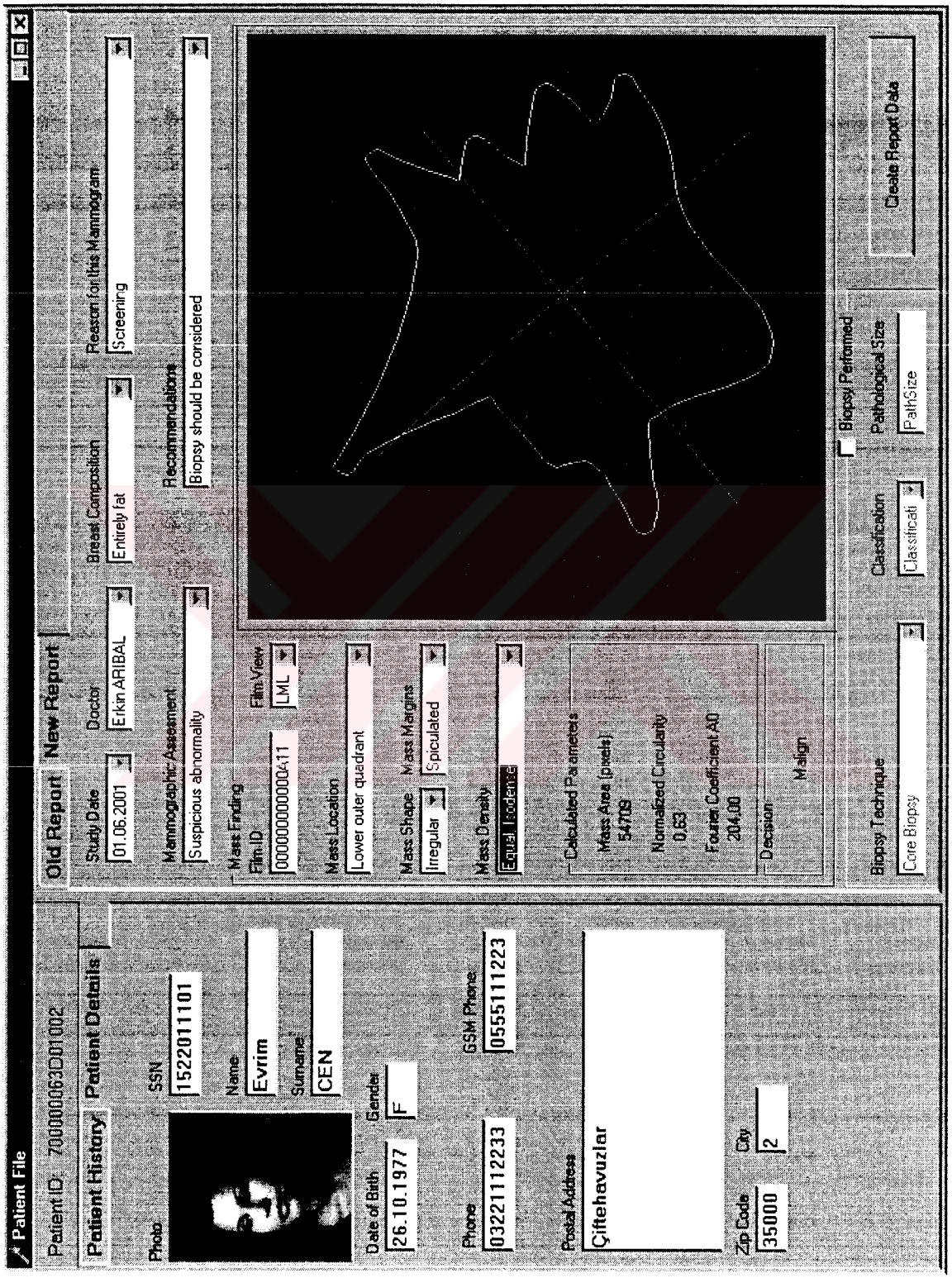

Figure 6.8 Patient file form after a new mass entry.

6.16 Patient Privacy

Touch Memory (TOM) buttons are electronic memory chips contained in small, water-resistant, stainless steel canisters. All Touch Memory buttons contain a unique ID number that is unalterable and identifies each button [47]. The unique ID number in the TOM is used as a tool that permits access to an electronic patient record in the mammogram database via the developed computer software. DS9097 COM port adapter is employed as a TOM reader and connected to the computer via an RS232 serial port.

7. AUTOMATIC PATIENT IDENTIFICATION

7.1 Introduction

Although human-readable labels have been used for ages, it was the advent of computer readable labels that quickly revolutionized the way drugstores operate and made possible the overnight delivery of drugs. When error prone and time consuming key entry was replaced by bar codes, it became convenient to build large databases to help in making accurate and timely decisions. In the next step in the evolution of labeling technology, ink on paper bar codes are surpasses by silicon media.

With automatic identification technology, a chip becomes the label that can serve as a standalone database. Attached to an object or carried by a person, the chip identifies and carries relevant information available instantly with little or no human intervention. People access secure areas with convenience and health care professionals accurately create records.

The lowest cost method of making a chip into a computer readable label is to reduce its many minute conductors to just one and extend it to a larger easy to contact point. The simplest arrangement is a single data conductors plus a ground contact. In this way a two piece stainless steel container called a microcan serves both as protective housing and electrical contacts: lid (data) and rim (ground). Its circular shape guides a simple, cup shaped probe over its rounded surfaces even if struck with significant misalignment. The 16 mm button shape serves all touch memories.

While touch memories share some of the characteristics of bar codes, these chip-based data carriers have many advantages over ink on paper technology:

Touch memories can be read without expensive electro-optical equipment.

Each touch memory proves its identity by its unique registration number. Steel container called microcan is better suited to harsh operating environments. All communication with touch memories is reduced to a single signal plus ground. Long and short pulses encode

the binary 1's and 0's. Because touch memories are digital circuits, they talk directly to other chips in a computer, resulting in minimal cost interface using one CMOS/TTL logic signal. A reader for touch memories can be implemented with just one spare I/O line of a microcomputer, often a free source in a system. The signaling rate of touch memory was chosen so that a computer's serial port operating at 115.2k bits per second can supply the timing by sending a bit for a byte.

Because of its unique serial number, low cost and reasonable access time for reading, touch memories are likely candidates for automatic patient identification. In the present work, the DS1990A iButton was used as the tool to permit access to the electronic patient record in the developed mammogram database.

DS9097A is a RS232C com port reader for iButtons. To enable DS1990A iButton readings via DS9097A, a dynamic link library for windows applications was written and standard procedure calls were employed in the software developments.

7.2 DS1990A Serial Number iButton

The DS1990A Serial Number iButton is a rugged data carrier that acts as an electronic registration number for automatic identification. The DS1990A consists of a factory–lasered, 64–bit ROM that includes an unique 48–bit serial number, an 8–bit CRC and an 8–bit Family Code (01h) as shown in Figure 7.1. Data is transferred serially via the 1–Wire protocol that requires only a single data lead and a ground return.

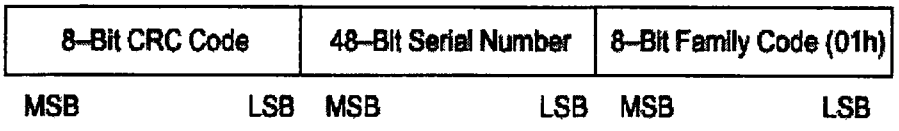

Figure 7.1 DS1990A memory map.

The durable microcan package is highly resistant to environmental hazards such as dirt, moisture and shock. Its compact coin–shaped profile, as shown in Figure 7.2, is self–aligning with mating receptacles, allowing the DS1990A to be used easily by patients. Accessories permit the DS1990A to be mounted on plastic key tabs, photo ID badges, printed circuit boards or any smooth surface of an object.

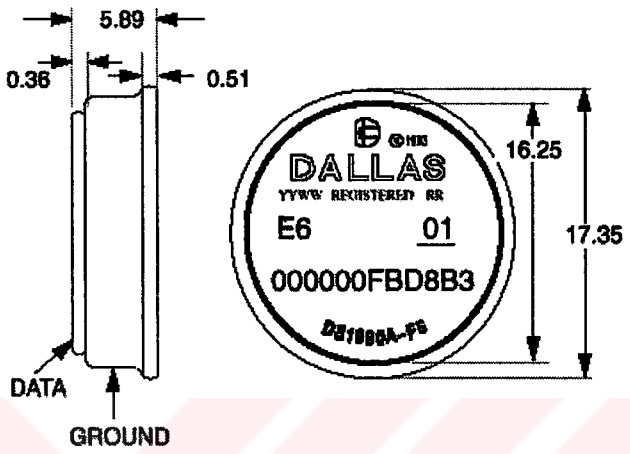

Figure 7.2 Physical dimensions of DS1990A F5.

The 1–Wire bus is a system, which has a single bus master system and one or more slaves. In all instances, the DS1990A is a slave device. The bus master is typically a microcontroller or a PC. The discussion of this bus system is broken down into three topics: hardware configuration, transaction sequence, and 1–Wire signaling (signal type and timing).

7.3 Hardware Configuration

The 1–Wire bus has only a single line by definition; it is important that each device on the bus be able to drive it at the appropriate time. To facilitate this, each device attached to the 1–Wire bus must have an open drain connection or 3–state outputs. The DS1990A is an open drain device. An internal equivalent circuit is given in Figure 7.3. The bus master can be the same equivalent circuit. If a bidirectional pin is not available, separate output and input pins can be tied together. The bus master requires a pull–up

resistor at the master end of the bus, with the standard TTL bus master circuit as shown in Figure 7.4.

The value of the pull–up resistor should be approximately 5 kΩ for short line lengths. A multidrop bus consists of a 1–Wire bus with multiple slaves attached. The 1–Wire bus has a maximum data rate of 16.3k bits per second.

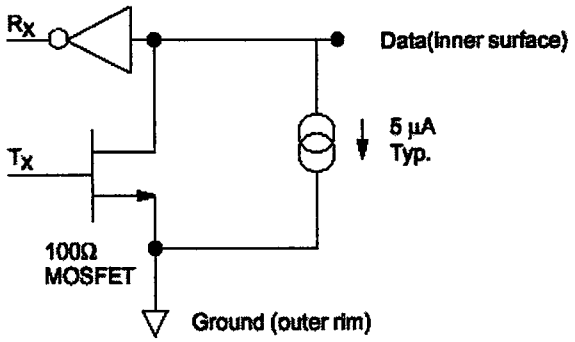

Figure 7.3 DS1990A equivalent circuit.

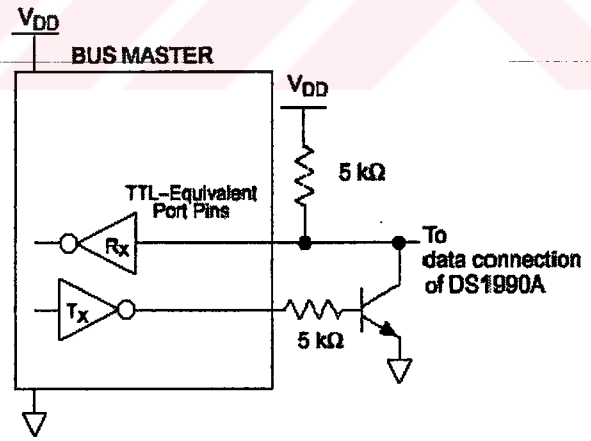

Figure 7.4 Standard TTL as bus master.

The idle state for the 1–Wire bus is high. If for any reason, a transaction needs to be suspended, the bus must be left in the idle state if the transaction is to resume. If this does

not occur and the bus is left low for more than 120 μs, one or more of the devices on the bus may be reset.

Transaction Sequence

The sequence for accessing the DS1990A via the 1–Wire port is as follows [47]:

Initialization: All transactions on the 1–Wire bus begin with an initialization sequence. The initialization sequence consists of a reset pulse transmitted by the bus master followed by a presence pulse(s) transmitted by the slave(s). The presence pulse lets the bus master know that the DS1990A is on the bus and is ready to operate.

ROM Function Command: Once the bus master has detected a presence, it can issue one of the four ROM function commands. All ROM function commands are eight bits long. Some of these commands are listed in Table 7.1.

Read Data

Table 7.1
Read and Search ROM Commands

Command	Description
Read ROM	This command allows the bus master to read the DS1990A' s 8-bit family code, unique 48-bit serial number, and 8-bit CRC. The DS1990A read ROM function will occur with a command type of either 33h or 0Fh.
Search ROM	When a system is initially brought up, the bus master might not know the number of devices on the 1–Wire bus or their 64–bit ROM codes. The search ROM command allows the bus master to use a process of elimination to identify the 64–bit ROM codes of all slave devices on the bus. This function will occur with a command type of F0h.

1-Wire Signaling

The DS1990A requires strict protocols to insure data integrity. The protocol consists of four types of signaling on one line: Reset sequence with reset pulse and presence pulse, write 0, write 1 and read data. All these signals except presence pulse are initiated by the bus master. The initialization sequence required to begin any communication with the DS1990A is shown in Figure 7.5. A Reset Pulse followed by a presence pulse indicates the DS1990A is ready to send or receive data given the correct ROM command.

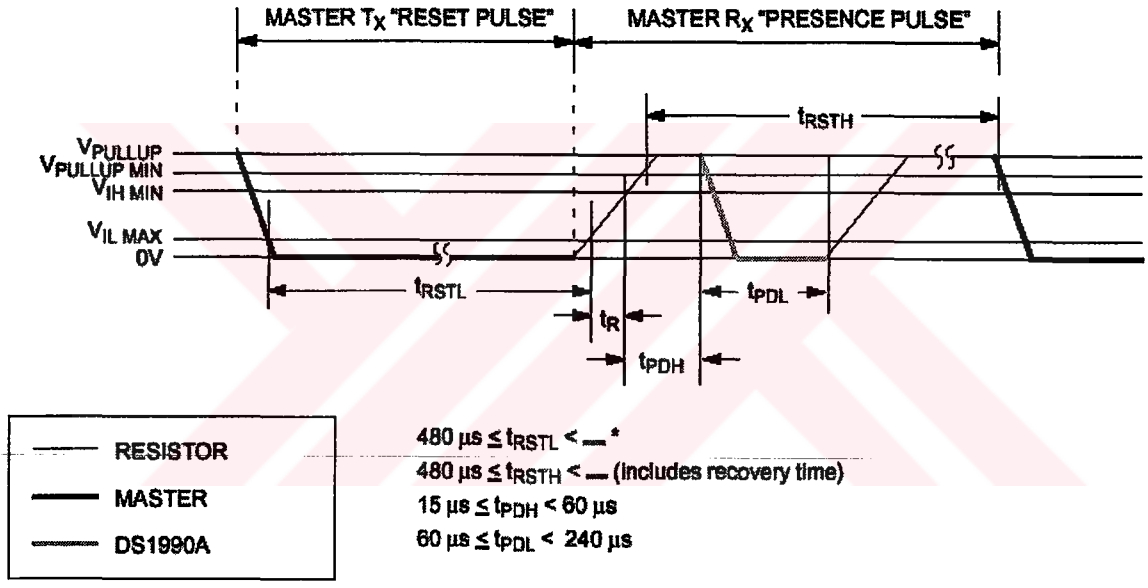

Figure 7.5 Initialization procedure, reset and presence pulses

The bus master transmits (TX) a reset pulse (a low signal for a minimum of 480 ms). The bus master then releases the line and goes into receive mode (RX). The 1–Wire bus is pulled to a high state via the 5 kΩ pull-resistor. After detecting the rising edge on the data contact, the DS1990A waits (t_{PDH} , 15-60 ms) and then transmits the presence pulse (t_{PDL} , 60-240 ms).

The definitions of write and read time slots are illustrated in Figures 7.6.a, 7.6.b and 7.6.c. The master driving the data line low initiates all time slots. The falling edge of the

data line synchronizes the DS1990A to the master by triggering a delay circuit in the DS1990A. During write time slots, the delay circuit determines when the DS1990A will sample the data line. For a read data time slot, if a "0" is to be transmitted, the delay circuit determines how long the DS1990A will hold the data line low overriding the 1 generated by the master. If the data bit is a "1", the iButton will leave the read data time slot unchanged.

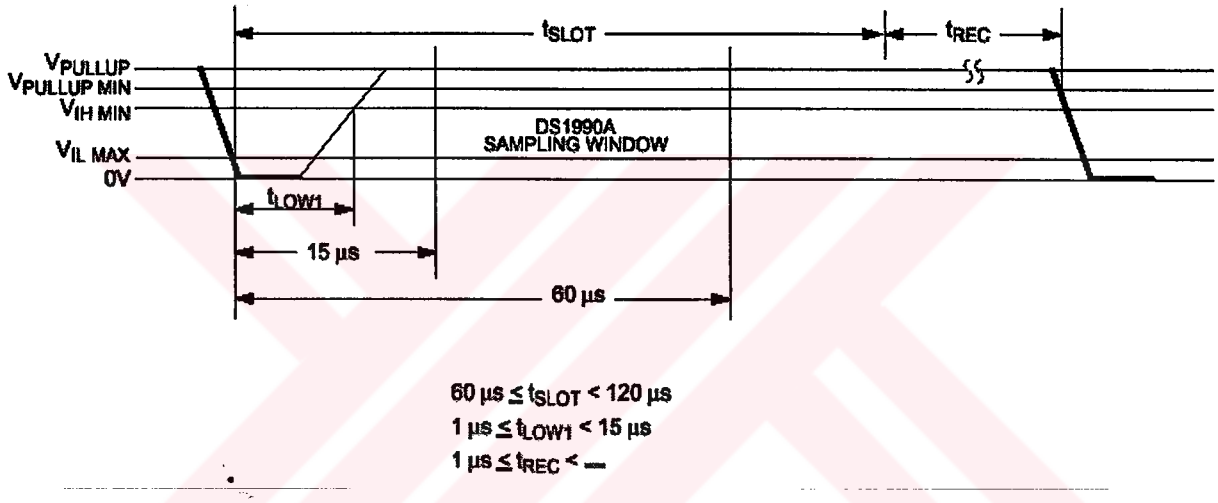

Figure 7.6.a Read / write timing diagram, write-one time slot.

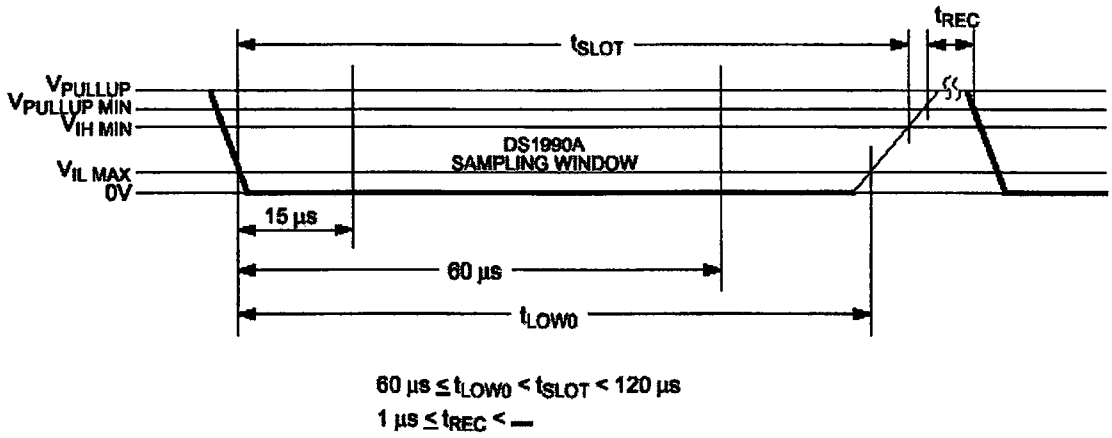

Figure 7.6.b Read / Write timing diagram, write-zero time slot.

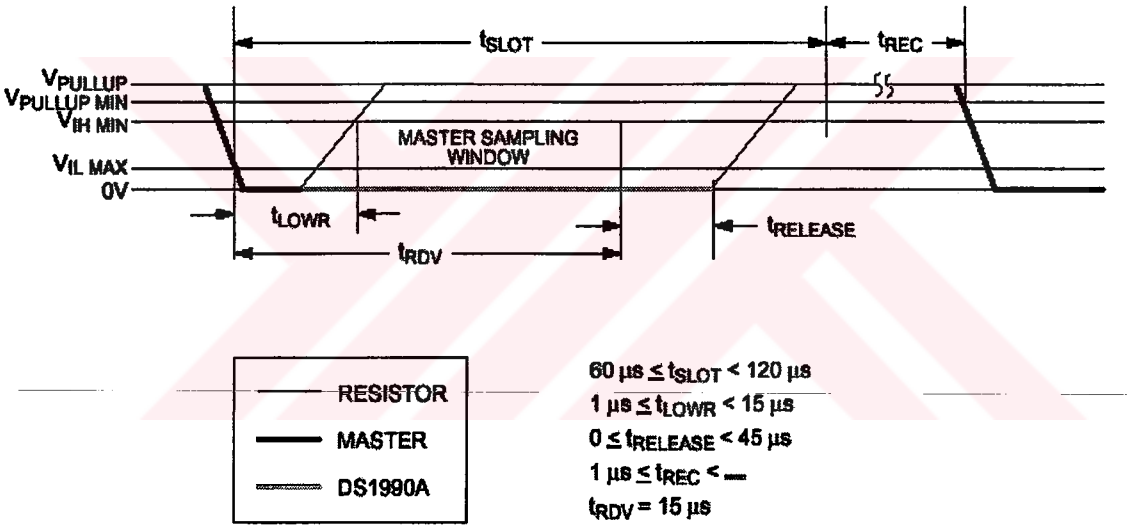

Figure 7.6.c Read / Write timing diagram, read .data time slot.

Cyclic Redundancy Check Generation

To validate the data transmitted from DS1990A, the bus master may generate a cyclic redundancy check value, called as CRC value, from the data as it is received. This generated value is compared to the value stored in the last eight bits of the DS1990A. The bus master computes the CRC over the 8 bit family code and all 48 ID number data bits but

not over the stored CRC value itself. If the two CRC values match, the transmission is error-free. The equivalent polynomial function of this CRC is given in (7.1).

$$CRC = x^8 + x^5 + x^4 + 1 \tag{7.1}$$

The 1-Wire CRC can be generated using a polynomial generator consisting of a shift register and XOR gates as shown in Figure 7.7.

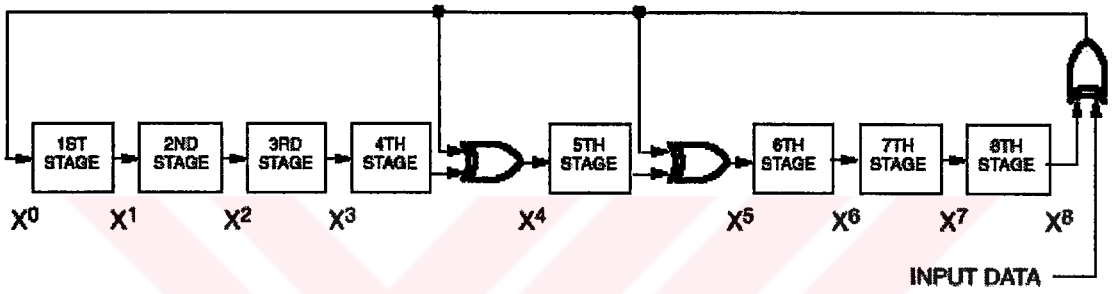

Figure 7.7 CRC generated using a polynomial generator consisting of a shift register and XOR gates.

The shift register bits are first initialized to zero. For the ROM section, starting with the least significant bit of the family code one bit at a time is shifted in. After the 8^{th} bit of the family code has been entered then the serial number is entered. After the 48^{th} bit of the serial number has been entered the shift register contains the CRC value. Shifting in the eight bits of the CRC should return the shift register to all zeros.

Interfaces

There exit two types of interfaces: TTL and RS232. TTL interface includes microprocessors and logic circuits; assign logic 0 up to 0.8V and logic 1 over 2.2V. In addition to this, voltages between –3V and –15V is logic 0, while voltages between +3V and +15V is logic 1 for RS232 interfaces. A typical interface circuit between TTL and 1-Wire system is shown in Figure 7.8. At this point one must remember that long distance

transmission lines will cause capacitive effects and delays. The value of the pull-up resistor in the circuit must be selected appropriately to detect logic levels.

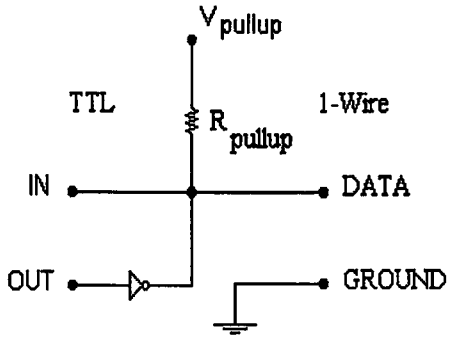

Figure 7.8 Interface between TTL and 1-wire system.

8250 UART is a typical controller for RS232 interface and undertakes the critical timing software execution load of microprocessor. The block diagram of 8250 UART is shown in Figure 7.9.

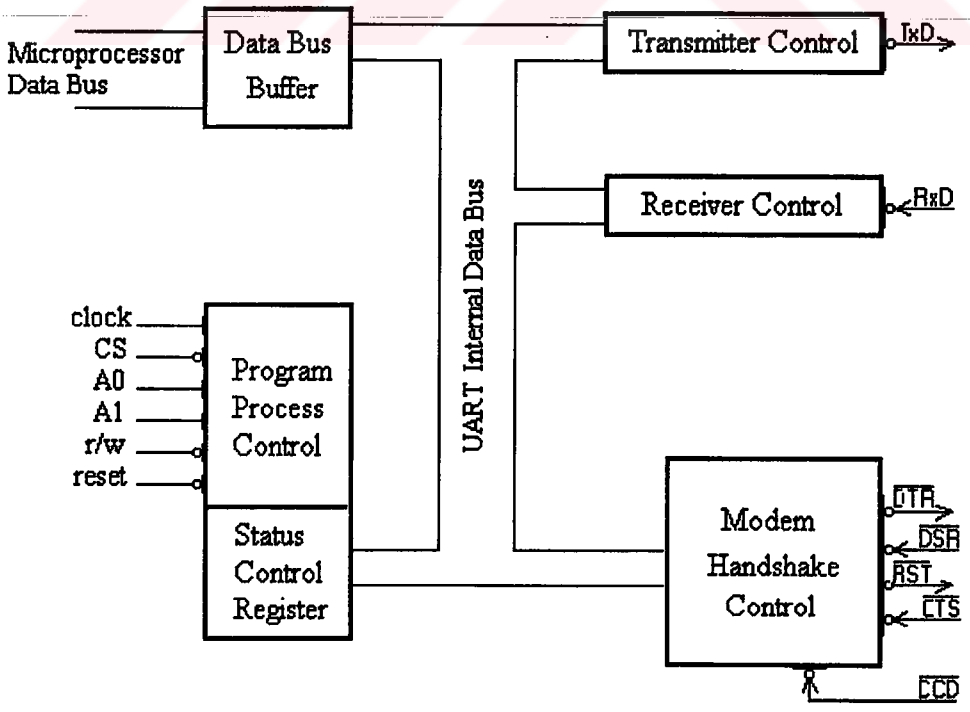

Figure 7.9 8250 UART block diagram.

For writing operation the microprocessor only puts the 8-bit data into the transmit register of UART, then UART executes the standard procedures while for reading operation the microprocessor reads the receive register of UART.

In all serial communications, to achieve proper run of UART, a software must use serial communication interrupts of the computer for setting the following parameters:

- Transmission speed,
- Data length,
- Number of parity bits,
- Number of start bits,
- Number of stop bits.

For 1-wire communication, the UART is set to a high transmission speed. Each received 8-bits represent a bit in 1-Wire bus system. This representation scheme must be applied during writing and reading processes. The easy way of this application can be achieved by employing a high level programming language such as Delphi.

The pin descriptions of 9 pin and 25 pin RS232 ports are listed in Table 7.2. While DTR (Data Terminal Ready) and RST (Request To Send) connections are required for 1-Wire system, there is no need to use other control connections.

Table 7.2
RS232 Connections for 9 and 25 Pins.

Signal	9-Pin Connect.	25-Pin Connect.	Description	Status
RXD	2	3	Receive Data	Input
TXD	3	2	Transmit Data	Output
DTR	4	20	Data Terminal Ready	Output
RTS	7	4	Request To Send	Output
GND	5	7	Ground	Reference
DSR	6	6	Data Set Ready	Input
CTS	8	5	Clear To Send	Input

7.4 DS9097A Com Port Adapter

The DS9097A COM Port Adapter, as shown in Figure 7.10, is a simple, low-cost passive adapter which performs RS232C (±12V) level conversion, allowing an iButton probe to be connected to the serial port of a computer so that it can read iButtons directly. The schematic layout of DS9097A is shown in Figure 7.11. The serial port must support a data transmission rate of 115.2 kbits/s in order to create the 1-Wire TM time slots correctly.

Nearly all PCs support the required data rate and are fully compatible with the DS9097A. Since an 8-bit character (6 data bits plus start and stop bit) on the RS232 port operating at 115.2 kbits/s is used to form a single 1-Wire time slot, the maximum effective 1-Wire transfer rate is 14.4 kbits/s (regular speed).

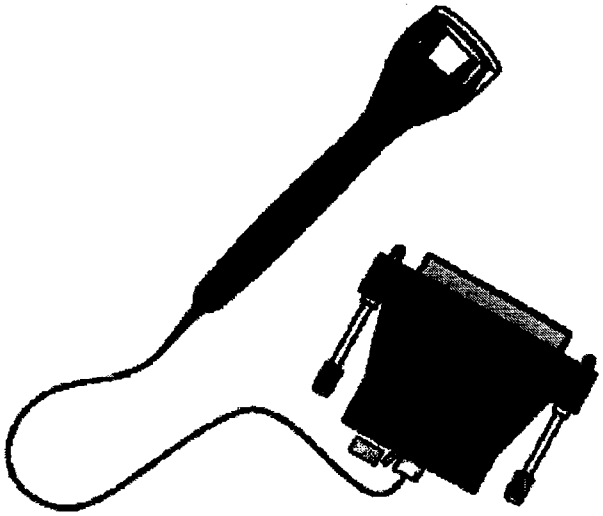

Figure 7.10 DS9097A com port reader as a iButton reader.

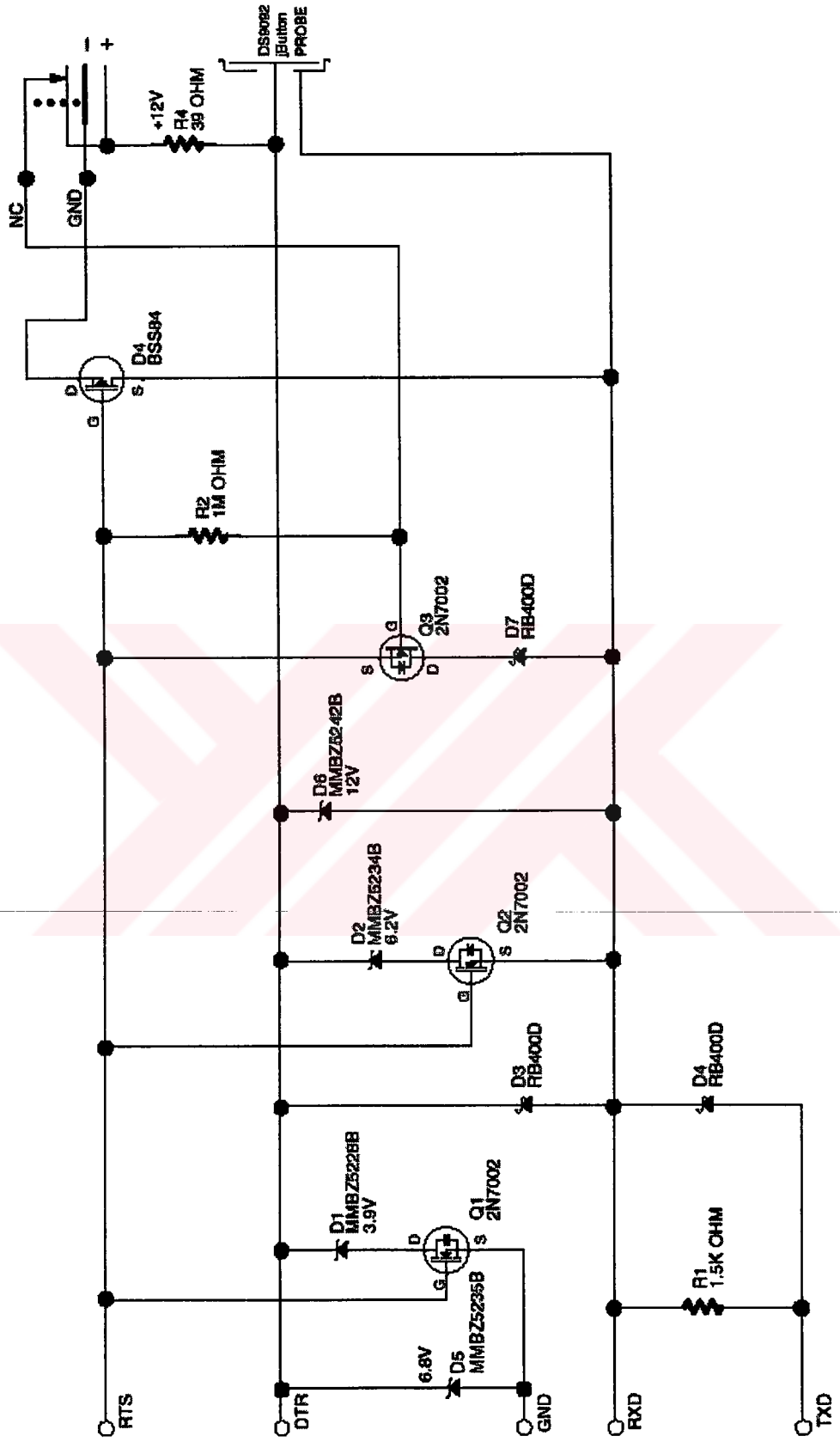

Figure 7.11 Schematic layout of DS9097A.

8. CONCLUSIONS

8.1 Introduction

Breast cancer is one of the most common malignancies in women and a rare malignancy in men. Women who are diagnosed at an early stage can survive this often deadly disease. Mammography provides the best screening modality for detecting early breast cancer, even before a lesion is palpable. Because of the malignant mass pathology, the shape of the mammographic mass can be used to discriminate between malignant and benign masses.

8.2 Results

In this study the use of shape features to classify breast masses has been investigated and a classification scheme has been developed to classify masses as either benign or malignant. To examine the classification scheme performance, a total of 25 biopsy-performed mammographic mass shapes have been analyzed by the software. The system successfully classified %90 of malignant masses and %93 of benign masses.

The mammogram database management system developed stores the images of the masses, calculated shape descriptor parameters and some additional data, such as patient history, category of the mass and biopsy report if performed which are required in BI-RADS. This database facilitates statistical data analysis and comparisons of previous and current parameters of the mammographic mass records obtained from the same patient.

A touch on memory (TOM) system has been used as a tool to gain easy and secure access to the electronic patient records in the mammogram database. Robustness, size, user-friendliness and cost considerations make the TOM system a wise choice. However, if the user decides to use other systems such as bar code or magnetic card systems, the

TOM reader can be replaced with a bar code or magnetic card reader without an extra software patch since they have been included in the present software.

The system has been developed for clinical use; it has a user-friendly graphical interface and runs on MS Windows operation systems. The software can make decisions in a few seconds and can be used as a teaching tool, a research tool and as a mammographer' s software assistant.

8.3 Recommendations for Future Work

In its present form, before the mammographic record is given to the developed system for processing, the contour of the suspicious mammographic mass must be drawn manually by an expert mammographer. This process, besides requiring an expert, takes-up some time for big masses. It is possible to automate this process by means an automatic mass boundary detection module. The design of this module requires a very large mammographic database and therefore was not included in the present work. In the future, this module can be developed using the database that will be build-up by the present system. When such a module is added the system can also be used for automated mammography screening.

In mammographic decisions, previous mammograms have proven to be very useful. The database developed facilitates decisions by providing fast and easy access to previous mammographic data and enabling rapid comparisons of the past and the present data. More sophisticated tools for further facilitating decisions can be developed using differential and dynamic analysis techniques. This task is left as an important future work. Another future work concerns the development of a module for automated detection of micro calcifications. These calcifications are also serious cause of breast malignancy and early detection for these calcifications is therefore also important.

APPENDIX A
THE ACR BREAST IMAGING REPORTING AND DATA SYSTEM (BI-RADS)

A.1 Introduction

The American College of Radiology (ACR) Breast Imaging Reporting and Data System, BI-RADS, is the product of a collaborative effort between members of various committees of the American College of Radiology with cooperation from the National Cancer Institute, the Centers for Disease Control and Prevention, the Food and Drug Administration, the American Medical Association, the American College of Surgeons, and the College of American Pathologists.

This system is a quality assurance tool designed to standardize mammographic reporting, reduce confusion in breast imaging interpretations, and facilitate outcome monitoring. Through a medical audit and outcome monitoring, the system provides important peer review and quality assurance data to improve the quality of patient care.

The ACR Breast Imaging Reporting and Data System is divided into four sections:

- Breast Imaging Lexicon
- Reporting System
- Follow-Up and Outcome Monitoring
- ACR National Mammography Database

A.2 Breast Imaging Lexicon

Terminology has evolved over many years, and the results have often led to confusion as to their meaning. The descriptive terms that follow are the terms and definitions that have been recommended by the ACR Task Force on Breast Cancer, and it is hoped they will be adopted by all those involved in breast imaging. It is believed that these terms provide a fairly complete categorization of lesions, but if there are any significant substantive changes, they may be submitted to the Task Force on Breast Cancer

significant substantive changes, they may be submitted to the Task Force on Breast Cancer of the American College of Radiology for review and inclusion if accepted by the Task Force.

Masses

A "MASS" is a space occupying lesion seen in two different projections. If a potential mass is seen in only a single projection it should be called a "DENSITY" until its three-dimensionality is confirmed.

(1) Shape

- Round: A mass that is spherical, ball-shaped, circular or globular.
- Oval: A mass that is elliptical or egg-shaped.
- Lobular: A mass that has contours with undulations.
- Irregular: The lesion's shape cannot be characterized by any of the above.

(2) Margins

- Circumscribed (Well-Defined or Sharply-Defined) Margins: The margins are sharply demarcated with an abrupt transition between the lesion and the surrounding tissue.
- Microlobulated Margins: The margins undulate with short cycles producing small undulations.
- Obscured Margins: One which is hidden by superimposed or adjacent normal tissue and cannot be assessed any further.
- Indistinct (Ill Defined) Margins: The poor definition of the margins raises concern that there may be infiltration by the lesion and this is not likely due to superimposed normal breast tissue.
- Spiculated Margins: The lesion is characterized by lines radiating from the margins of a mass.

(3) Density

This is used to define the x-ray attenuation of the lesion relative to the expected attenuation of an equal volume of fibroglandular breast tissue. It is important in that most breast cancers that form a visible mass are of equal or higher density than an equal volume of fibroglandular tissue. It is rare (although not impossible) for breast cancer to be lower

in density. Breast cancers are never fat containing (radiolucent) although they may trap fat.

- High density
- Equal density (isodense)
- Low density (lower attenuation, but not fat containing)
- Fat containing - radiolucent. This includes all lesions containing fat such as an oil cyst, lipoma, or galactocele as well as mixed lesions such as the hamartoma or fibroadenolipoma. [When appropriate, histologic terms may be included]

Calcifications

Benign calcifications are usually larger than calcifications associated with malignancy. They are usually coarser, often round with smooth margins and are much more easily seen. Calcifications associated with malignancy are usually very small and often require the use of a magnifying glass to see them well.

When a specific etiology cannot be given, a description of calcifications should include the morphology and distribution of the calcifications. Benign calcifications need not always be reported. They should be reported if the interpreting radiologist is concerned that they might be misinterpreted by other observers. Types and distribution of calcifications are as follows:

Typically Benign -

- Skin Calcifications: These are typical lucent centered deposits that are pathognomonic. Atypical forms may be confirmed by tangential views to be in the skin.
- Vascular Calcifications: Parallel tracks, or linear tubular calcifications that are clearly associated with blood vessels.
- Coarse or ("Popcorn Like") Calcifications: These are the classic calcifications produced by an involuting fibroadenoma.
- Large Rod-Like Calcifications: These are benign calcifications forming continuous rods that may occasionally be branching, are usually more than 1 mm in diameter, may have lucent centers, if calcium surrounds rather than fills an ectactic duct.

These are the kinds of calcifications found in secretory disease, "plasma cell mastitis", and duct ectasia.

- Round Calcifications: When multiple, they may vary in size. They are usually considered benign and when small [under 1 mm], they frequently are formed in the acini of lobules. When under 0.5 mm the term punctate can be used.
- Lucent-Centered Calcifications: These are benign calcifications that range from under 1 mm to over a centimeter or more. These deposits have a smooth surfaces, are round or oval, and have a lucent center. The "wall" that is created is thicker than the "rim or eggshell" type of calcifications. Included are areas of fat necrosis, calcified debris in ducts, and occasional fibroadenomas.
- Eggshell or Rim Calcifications: These are very thin benign calcifications that appear as calcium deposited on the surface of a sphere. These deposits are usually under 1 mm in thickness when viewed on edge. Although fat necrosis can produce these thin deposits, calcifications in the wall of cysts are the most common "rim" calcifications.
- Milk of Calcium Calcifications: This is consistent with sedimented calcifications in cysts. On the craniocaudal image they are often less evident and appear as fuzzy, round, amorphous deposits while on the 90° lateral, they are sharply defined, semilunar, crescent shaped, curvilinear (concave up), or linear defining the dependent portion of cysts.
- Suture Calcifications: These represent calcium deposited on suture material. These are relatively common in the post-irradiated breast. They are typically linear or tubular in appearance and knots are frequently visible.
- Dystrophic Calcifications: These are calcifications that usually form in the irradiated breast or in the breast following trauma. Although irregular in shape, they are usually over 0.5 mm in size. They often have lucent centers.
- Punctate Calcifications: These are round or oval, less than 0.5 mm with well-defined margins.

Intermediate Concern Calcifications -

- Amorphous or Indistinct Calcifications: These are often round or "flake" shaped calcifications that are sufficiently small or hazy in appearance that a more specific morphologic classification cannot be determined.

Higher Probability Of Malignancy -

- Pleomorphic or Heterogeneous Calcifications (Granular): These are usually more conspicuous than the amorphic forms and are neither typically benign (see above) nor typically malignant (see below) irregular calcifications with varying sizes and shapes that are usually less than 0.5 mm in diameter.
- Fine, Linear or Fine, Linear, Branching (Casting) Calcifications: These are thin, irregular calcifications that appear linear, but are discontinuous and under 0.5 mm in width. Their appearance suggests filling of the lumen of a duct involved irregularly by breast cancer.

Distribution Modifiers

These are used as modifiers of the basic morphologic description and describe the arrangement of the calcifications. Multiple similar groups may be indicated when there is more than one group of calcifications that are similar in morphology and distribution.

- Grouped or Clustered [Although historically the term "clustered" has connoted suspicion, the term shall now be used as a neutral distribution modifier and may reflect benign or malignant processes]: Should be used when multiple calcifications occupy a small volume [less than 2 cc] of tissue.
- Linear: Calcifications arrayed in a line that may have branch points.
- Segmental: These are worrisome in that their distribution suggests deposits in a duct and its branches raising the possibility of multifocal breast cancer in a lobe or segment of the breast. Although benign causes of segmental calcifications exist such as "secretory disease" this distribution is of greater concern when the morphology of the calcifications is not specifically benign.
- Regional: These are calcifications scattered in a large volume of breast tissue not necessarily conforming to a duct distribution that are likely benign, but are not everywhere in the breast, and do not fit the other more suspicious categories.
- Diffuse/Scattered: These are calcifications that are distributed randomly throughout the breast.

Multiple similar groups may be indicated when there is more than one group of calcifications that are similar in morphology and distribution.

Architectural Distortion

The normal architecture is distorted with no definite mass visible. This includes spiculations radiating from a point, and focal retraction or distortion of the edge of the parenchyma. Architectural distortion can also be an associated finding.

Special cases

Tubular Density/Solitary Dilated Duct: This is a tubular or branching structure that likely represents a dilated or otherwise enlarged duct. If unassociated with other suspicious clinical or mammographic findings it is usually of minor significance.

- Intramammary Lymph Node: These are typically reniform or have a radiolucent notch due to fat at the hilum and are generally 1 cm. or smaller in size. They may be larger than 1 cm. and normal when fat replacement is pronounced. They may be multiple, or marked fat replacement may cause a single lymph node to look like several rounded masses. This specific diagnosis should be made only for masses in the lateral half and usually upper portion of the breast, although on rare occasions they may be in other areas of the breast.
- Asymmetric Breast Tissue: Asymmetric breast tissue is judged relative to the corresponding area in the other breast and includes a greater volume of breast tissue, greater density of breast tissue, or more "prominent ducts." There is no focal mass formation, no central density, no distorted architecture, and no associated calcifications. Asymmetric breast tissue usually represents a normal variation, but may be significant when it corresponds to a palpable asymmetry.
- Focal Asymmetric Density: This is a density that cannot be accurately described using the other shapes. It is visible as asymmetry of tissue density with similar shape on two views, but completely lacking borders and the conspicuity of a true mass. It could represent an island of normal breast, but its lack of specific benign characteristics may warrant further evaluation. Additional imaging may reveal a true mass or significant architectural distortion.

Associated Findings

When no other abnormality is present, the following findings are used with masses or calcifications or may stand alone as FINDINGS;

- Skin Retraction: The skin is pulled in abnormally.
- Nipple Retraction: The nipple is pulled in or inverted.
- Skin Thickening: This may be focal or diffuse.
- Trabecular Thickening: This is a thickening of the fibrous septae of the breast.
- Skin Lesion: Commented on when it projects over the breast in two views and may be mistaken for an intramammary lesion.
- Axillary Adenopathy: Enlarged non-fatty replaced axillary lymph nodes may be commented on. Mammographic assessment of these nodes is unreliable.
- Architectural Distortion: As an ASSOCIATED FINDING it can be used in conjunction with a FINDING to indicate that the normal tissue structure is distorted or retracted surrounding the FINDING.
- Calcifications: As an ASSOCIATED FINDING it can be used in conjunction with a FINDING to describe calcifications within or immediately adjacent to the FINDING.

Location of Lesion

A significant lesion must always be triangulated so that its three-dimensional location within the breast is known. This usually requires it to be visible on two mammographic projections. This is more precise if the lesion is visible on orthogonal views. The location of the lesion should be described using the clinical orientation extrapolated from the film location. The breast is viewed as the face of a clock with the patient facing the observer. Use of quadrants to describe location is an option. Use of both clockface and quadrant is encouraged. The side is given first, followed by the location and depth of the lesion. Depth divides the breast arbitrarily into anterior, middle and posterior thirds. Immediately beneath the nipple is the subareolar region. For locations use clockface preceded by left or right or both for side such as upper outer quadrant, upper inner quadrant, lower outer quadrant, and lower inner quadrant or use subareolar, central, and axillary tail. Subareolar, axillary tail, and central do not require depth. Subareolar, central, and axillary tail do not require clockface location. For depth add Anterior, Middle, and Posterior.

A.3 Reporting System

The reporting system is designed to provide an organized approach to image interpretation and reporting. It does not require a computer system, but the utilization of a computer in reporting is strongly encouraged. Not only does this facilitate reporting, but data are simultaneously collected for the maintenance of the recommended database for future review. This will permit individual radiologists or groups to monitor their own results and appraise accuracy in image interpretation, and adjust thresholds appropriately. There is no ideal computer system, but it is strongly recommended that the system used require a minimum of interaction. The radiologist's attention should be focused on the interpretation of the images. The simplest input utilizes a single screen with minimal interaction needed from the radiologist. The goal is to maximize the image viewing time, and minimize any distractions from the reporting.

Report Organization

Use of approved terminology is encouraged. This system categorizes the overall composition of the breast and then describes lesions by their basic geometry, border characteristics, and density. Calcifications in the system are described according to size, morphology, and distribution. The findings are then interpreted and an assessment rendered that includes the degree of concern, and any pertinent recommendations. Thus, the breast imaging report should be divided into breast composition, findings and overall assessment.

Breast Composition

This is an overall assessment of the attenuating tissues in the breast to help indicate the relative possibility that a lesion could be hidden by the normal tissues. Generally, this includes fatty, mixed or dense. Since mammography cannot detect all breast cancers, physical examination is always a key element of screening. It is important to alert the clinician that in the radiographically dense breast the ability of mammography to detect small cancers is reduced. Although mammography is still useful in these women, the physical examination (which is always important) is increased in importance. The available data do not support the use of mammographic patterns for determining screening frequency (i.e., risk for breast cancer) If an implant is present, it should be stated in the

report and an implant description code added as appropriate. For consistency, breast composition should be included for all patients using the following patterns:

- The breast is almost entirely fat.
- There are scattered fibroglandular densities.
- The breast tissue is heterogeneously dense. This may lower the sensitivity of mammography.
- The breast tissue is extremely dense, which could obscure a lesion on mammography.

Findings

(1) A clear description of any significant finding.

- Mass:

  Size

  Lesion type and modifiers

  Associated calcifications

  Associated findings

  Location

  *How changed, if previously present.

- Calcifications:

  Morphology - type or shape and modifiers

  Distribution

  Associated findings

  Location

  *How changed, if previously present.

- Architectural Distortion:

  Associated calcifications

  Associated findings

  Location

*How changed, if previously present.

- Special Cases:

  Associated calcifications

  Associated findings

  Location

  *How changed, if previously present.

The clinical location of the abnormality as extrapolated from the mammographic location (based on the face of a clock and/or quadrant).

(2) An overall (summary) impression:

All final impressions should be complete with each lesion fully categorized and qualified. An indeterminate reading should only be given in the screening setting where additional imaging evaluation is recommended before a final opinion can be rendered.

Interpretation is facilitated by recognizing that most mammograms can be categorized under a few headings. These are listed below, and suggested codes are included for computer use. If a suspicious abnormality is detected, the report should indicate that biopsy should be considered. This is an assessment where the radiologist has sufficient concern that biopsy is warranted unless there are other reasons why the patient and her physician might wish to defer the biopsy.

Assessment Categories

(1) Assessment Is Incomplete

Category 0: Need Additional Imaging Evaluation

Finding for which additional imaging evaluation is needed. This is almost always used in a screening situation and should rarely be used after a full imaging work up. A recommendation for additional imaging evaluation includes the use of spot compression, magnification, special mammographic views, ultrasound, etc. Whenever possible, the present mammogram should be compared to previous studies. The radiologist should use judgment in how vigorously to pursue previous studies.

(2) Assessment Is Complete - Final Categories

Category 1: Negative

There is nothing to comment on. The breasts are symmetrical and no masses, architectural disturbances or suspicious calcifications are present.

Category 2: Benign Finding

This is also a negative mammogram, but the interpreter may wish to describe a finding. Involuting, calcified fibroadenomas, multiple secretory calcifications, fat containing lesions such as oil cysts, lipomas, galactoceles, and mixed density hamartomas all have characteristic appearances, and may be labeled with confidence. The interpreter might wish to describe intramammary lymph nodes, implants, etc. while still concluding that there is no mammographic evidence of malignancy.

Category 3: Probably Benign Finding - Short Interval Follow-Up Suggested:

A finding placed in this category should have a very high probability of being benign. It is not expected to change over the follow-up interval, but the radiologist would prefer to establish its stability. Data are becoming available that shed light on the efficacy of short interval follow-up. At the present time, most approaches are intuitive. These will likely undergo future modification as more data accrue as to the validity of an approach, the interval required, and the type of findings that should be followed.

Category 4: Suspicious Abnormality - Biopsy Should Be Considered

These are lesions that do not have the characteristic morphologies of breast cancer but have a definite probability of being malignant. The radiologist has sufficient concern to urge a biopsy. If possible, the relevant probabilities should be cited so that the patient and her physician can make the decision on the ultimate course of action.

Category 5: Highly Suggestive of Malignancy - Appropriate Action Should Be Taken

These lesions have a high probability of being cancer.

Wording The Report

When available, the present examination should be compared to previous studies, and this should be indicated in the report. Reports should be organized with a brief description of the composition of the breast, any pertinent FINDINGS, followed by the ASSESSMENT with any recommendations. The report should be succinct using terminology from the approved lexicon without embellishment. Definitions and descriptors of the lexicon terms do not appear in the report narrative. Following the impression section of the report, both the assessment category number and the lexicon terminology for the assessment category should be stated.

A.4 Follow-Up And Outcome Monitoring

This section on the mammography audit describes certain minimum data to be collected and utilized to calculate important derived data which allow each radiologist to assess his or her overall performance in mammography interpretation. In addition to the basic clinically relevant audit, more complete mammography audit data may also be collected and utilized to calculate derived data to provide other important information regarding mammographic performance. To make sure that these data are protected as peer review information, radiologists should consult applicable state law and regulations.

A.5 ACR National Mammography Database

The maintenance of a database is an important quality assurance element of the ACR BI-RADSTM. Without monitoring the results of screening, it is impossible to know the success of the program. Each group should maintain the suggested data so that the accuracy of the individual screening programs and their success in diagnosing earlier stage breast cancers can be determined. This will allow each group to adjust its thresholds by comparison with pooled national data.

REFERENCES

1. Kopans, D.B., *Breast Imaging*, Lippincott-Raven Publishers, Philadelphia, 1998.
2. Kopans, D.B., J.E. Meyer, and N. Sadowsky, "Breast Imaging," *N Eng. J Med.*, Vol. 310, pp. 960, 1984.
3. Kopans, D.B., "Breast Imaging and The Standard of Care for the Symptomatic Patient," *Radiology*, Vol. 187, pp. 608-611, 1993.
4. Meyer, J.E., and D.B. Kopans, "Breast Physical Examination by The Mammographer: An Aid to Improved Diagnostic Accuracy," *Applied Radiology*, pp. 103-106, 1983.
5. Faulk, R.M. and E.A. Sickles, "Efficacy of Spot Compression Magnification and Tangential Views in Mammographic Evaluation of Palpable Masses," *Radiology*, Vol. 185, pp. 87-90, 1992.
6. Kopans, D.B., "Mammographic Screening for Breast Cancer," *Cancer*, Vol. 72, pp. 1809-1812, 1993.
7. King, M.J., "Mammography Screening for Breast Cancer," *Cancer*, Vol. 73, pp. 2003-2004, 1994.
8. Maass, H., "Mammakarzionom: Epideiologie," *Gynakologe*, Vol. 27, pp. 3, 1994.
9. Stoll, B.A., *Approaches to Breast Cancer Prevention*, Kluwer, London, 1991..
10. Friedrichs, K., "Genetische Aspekte des Mammakarzinoms," *Gynakologe*, Vol. 27, pp. 7, 1994.
11. Dupont, W.D., and D.L. Page, "Risk Factors for Breast Cancer in Women with Proliferative disease," *N Engl J Med.*, Vol. 312, pp. 146, 1985.
12. Heywang-Köbrunner, S., I. Schreer, and D. Dershaw, *Diagnostic Breast Imaging*, Thieme Stuttgart, New York, 1997.
13. Christensen, E.E., T.S. Curry, and J.E. Dowdey,, *An Introduction to the Physics of Diagnostic Radiology*, Lea Febinger, Philadelphia, 1978.
14. Kimme-Smith, C., L.W. Bassett, and R.H. Gold, *Workbook for Quality Mammography*, Williams & Wilkins, Baltimore, 1992.
15. Barnes, G.T., and I.A. Brezovich, "Contrast: Effect of Scattered radiation," in Logan WW, Muntz EP eds. Reduced Dose Mammography, Masson, New York, 1979..
16. Bird, R.E., L.W. Wallance, and B.C. Yankaskas, "Analysis of Cancers Missed at Screening Mammography," Radiology, Vol. 184, pp. 613, 1992.
17. Van Dick, J.A.A.M., A.L.M. Verbeek, J.H.C.L. Hendricks, and R. Holland, "The Current Detectability of Breast Cancer in a Mammographic Screening Program," *Cancer*, Vol. 72, pp. 1933, 1993.
18. Sickles, E.A., S.H. Ominski, and R.A. Sollitto, "Medical Audit of a Paired Throughput Mammography Screening Practice: Methodology and Results of 27114 Examinations," *Radiology*, Vol. 175, pp. 323, 1990.
19. Homer, M.J., "Breast Imaging : Pitfalls, Controversies and Some Practical Thoughts," *Radiol. Clin. North Am*, Vol. 23, pp. 459, 1985.
20. Page, D.L., and A.C. Winfield, "The Dense Mammogram," *American Journal of Radiology*, Vol. 147, pp. 487, 1986.
21. D'Orsi, C.J., "To Follow or Not To Follow, That is The Question," *Radiology*, Vol. 184, pp. 306, 1992.
22. Knutten, A.M., and J.J. Gisvold, "Likelihood of Malignant Disease for Various Categories of Mammographically Detected Nonpalpable Breast Lesions," *Radiology*, Vol. 189, pp. 927, 1993.

23. Rubin, E., "Critical Pathways in The Analysis of Breast Masses," *Radiographics*, Vol. 15, pp. 925, 1995.

24. Tabar, L., and P.B. Dean, "Optimum Mammography Technique: The annotate Cookbook Approach," *Admin Radiol.*, 54, 1989.

25. Eklund, G.W., and G. Cardenosa, " The Art of Mammographic Positioning," *Radiol Clin Nort Am.*, Vol. 30, pp. 21, 1992.

26. Feig, S.A., and R. McLelland, *Breast Carcinoma: Current Diagnosis and Treatment*, Masson, New York, 1983.

27. Martin, J.E., and H.S. Gallager, "Mammographic Diagnosis of Minimal Breast Cancer," *Cancer*, vol. 23, pp. 855-873, 1971.

28. Pratt, W.K., *Digital Image Processing*, Wiley, New York, 1991.

29. Harley, R.M., and A.R. Weeks, *Computer Imaging Recipes in C*, Prentice Hall, New York, 1993.

30. Pitas, I., *Digital Image Processing Algorithms*, Prentice Hall, New York, 1993.

31. Hu, M.K., "Visual Pattern Recognition by Moment Invariants" *IRE Trans. Info. Theory*, Vol. IT-8, pp: 179-187, 1962.

32. Gonzales, R.C., and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

33. Shen, L., R. M. Rangayyan, and J.E.L. Desautels, "Application of shape analysis to mammographic calcifications," IEEE Trans. Med. Imag., vol. 12, pp. 263-274, 1994.

34. Luo, X.M., "An Expert System for Mammogram Analysis", M.S. Thesis, University of Tennessee, Knoxville, TN, 1989.

35. Brzakovic, D., X.M. Luo, and P. Brzakovic, "An approach to automated detection of tumors in mammograms," *IEEE Trans. Med. Imag.*, Vol. 9, pp. 233-241, 1990.

36. Kilday, J., F. Palmieri, and M.D. Fox, "Classifying mammographic lesions using computerized image analysis," *IEEE Trans. Med. Imag.*, Vol. 12, pp. 664-669, 1993.

37. Melsa, L.J., and D.L. Cohn, *Decision and Estimation Theory*, Mc-Graw Hill, New York, 1978.

38. Metz, C.E., "Basic Principles of ROC Analysis", *Sem Nuc Med.*, Vol. 8, pp. 283-298, 1978.

39. Metz, C.E., "ROC Methodology in Radiological Imaging", *Invest Radiol.*, Vol. 21, pp. 720-733, 1986.

40. Griner, P.F., R.J. Mayewski, A.I. Mushlin, and P. Greenland, "Selection and Interpretation of Diagnostic Tests and Procedures", *Annals of Internal Medicine,* Vol. 94, pp. 555-600, 1981.

41. Zweig, M.H, and G. Campbell, "Receiver-operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine", *Clinical Chemistry,* Vol. 39, pp. 561-577, 1993.

42. Swets J.A., "ROC Analysis Applied to The Evaluation of Medical Imaging Techniques", *Invest Radiol,* Vol. 4, pp. 109-121, 1979.

43. Hanley, J.A., and B.J. McNeil, "The Meaning and Use of The Area Under a Receiver Operating Characteristic Curve", *Radiology*, Vol. 143, pp. 29-36, 1982.

44. Hanley, J.A., and B.J. McNeil, "A Method of Comparing The Areas Under Receiver Operating Characteristic Curves Derived from The Same Cases", *Radiology,* Vol. 148, pp. 839-843, 1983.

45. McClish, D.K., "Comparing The Areas Under More Than Two Independent ROC Curves", *Medical Decision Making*, Vol. 7, pp. 149-155, 1987.

46. Stephens, R.K., R.R. Plew, B. Morgan, and J. Perkins, *Teach Yourself SQL in 21 Days*, Macmillan Computer Publishing, New York, 1997.

47. *Automatic Identification Data Book*, Dallas Semiconductor Co., 1996.