2003

Computer assisted screening of digital mammogram images

John Terry Sample
Louisiana State University and Agricultural and Mechanical College

COMPUTER ASSISTED SCREENING OF DIGITAL MAMMOGRAM IMAGES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Computer Science

by
John Terry Sample
B.S., University of Southern Mississippi, 1999
August 2003

Acknowledgements

I would like to thank my advisor, Dr. John Tyler, for his support, encouragement, and patience during my years of graduate study. His experience and knowledge made this research possible.

I also thank my committee members, Dr. Young H. Chun, Dr. S.S. Iyengar, Dr. Donald Kraft, and Dr. Erno Sajo. They provided an essential review of my research and gave valuable guidance.

My colleagues at the Naval Research Laboratory, Frank McCreedy, Hillary Mesick, and Patrick McDowell have given help and advice in completion of this dissertation.

Finally, I thank Monica Azzolini. She provided invaluable help in preparing this dissertation.

Table of Contents

Acknowledgements iii

Abstract vii

Chapter 1 Introduction 1
1.1 Objectives of the Dissertation 1
1.2 Contributions of the Dissertation 3
1.3 Overview of Dissertation 3

Chapter 2 Breast Cancer 6
2.1 Breast Cancer 6
2.2 Screening Mammography 9
2.2.1 The Mammography Process 9
2.2.2 Mammographic View of Breast Anatomy 10
2.2.3 Objects in Mammogram Images 12
2.2.4 Background Observations in Mammogram Images 20
2.3 Mammogram Film Reading 21
2.3.1 Reading Errors 21
2.3.2 Prompting Systems 22

Chapter 3 Image Fundementals 23
3.1 Digital Mammograms as Images 23
3.1.1 Digital Images 23
3.2 Image Statistics 24
3.3 Image Enhancement 32
3.3.1 Histogram Operations 32
3.3.2 Filtering and Convolution 34
3.3.3 Image Smoothing 35
3.3.4 Image Sharpening 36
3.4 Image Segmentation 38
3.4.1 Edge Based Image Segmentation 39
3.4.2 Region Based Image Segmentation 40
3.5 Image Formats and Storage Methods 41

Chapter 4 Digital Mammography and Computer Aided Screening 42
4.1 Digital Mammogram Images 42
4.2 Preprocessing Methods 43
4.2.1 Image Enhancement 43
4.2.2 Image Segmentation 45
4.3 Abnormality Detection Techniques 47
4.3.1 Microcalcification Detection 47
4.3.2 Mass Detection 48

Chapter 5 Segmenting the Mammogram Image 51
5.1 Importance of Proper Image Segmentation 51
5.2 The Image Segmentation Algorithm 52
5.2.1 Histogram Based Segmentation 54
5.2.2 The Histogram Difference Method 55
5.2.3 Finding Component Boundaries 62
5.2.4 Adding Spatial Context 67
5.2.5 Segmenting Dense Mammogram Images 71
5.3 Process Summary 72

Chapter 6 Detecting Mass in Mammogram Images 73
6.1 Algorithms for Locating Masses in Digital Mammograms 73
6.2 Mass Detection Algorithm 74
6.2.1 A Template for Mammogram Masses 74
6.2.2 Template Matching Methods 79
6.2.3 Clustering and Segmentation of Suspicious Areas 84
6.2.4 Multi-Scale Enhancements 86
6.3 Process Summary 88

Chapter 7 Screening Algorithm Testing System 89
7.1 Results of Image Filters 90
7.2 Clustering Pixel Groups 91
7.3 Comparing to Pathology 101
7.4 Process Parameters 103
7.5 Results Recording 105
7.6 Process Summary 106

Chapter 8 Experimental Results 107
8.1 Developmental Image Sets 107
8.2 Segmentation Algorithm Testing 108
8.2.1 Experimental Methodology 108
8.2.2 Results 109
8.3 Testing of Mass Detection Algorithm 110
8.3.1 Experimental Methodology 110
8.3.2 Free-Response Receiver Operating Characteristic Curve 111
8.3.3 Results for Initial Mass Location 112
8.3.4 Results for Multiscale Mass Location 115
8.3.5 Results with Multiscale Templates at Initial Examination 117
8.3.6 Results with Limiting the Number of Detections 118
8.4 Discussion 121

Chapter 9 Summary 122
9.1 Review 122
9.2 Future Research 123
9.3 Conclusion 123

References 124

Vita 127

Abstract

The use of computer systems to assist clinicians in digital mammography image screening has advantages over traditional methods. Computer algorithms can enhance the appearance of the images and highlight suspicious areas. Screening provides a more thorough examination of the images. Any computer system that does screening of digital mammograms contains components to address multiple tasks such as: image segmentation, mass lesion detection and classification, and microcalcification detection and classification.

This dissertation provides both effective and efficient improvements to existing algorithms, which segment mammogram images and locate mass lesions. In addition, we provide a new algorithm to evaluate and report the results for mass lesion detection.

The algorithm presented for mammogram segmentation uses a histogram based operator to define the boundaries between the different components of a mammogram image. It employs a unique clustering algorithm to produce closed, labeled sets of pixels which represent the distinct image components.

The mass location algorithm uses a variation of template matching to locate suspicious areas. An evaluation of potential templates and algorithms is included. The method for testing and recording the results of the mass location algorithm groups suspicious pixels into regions and then compares them to the pathology.

Chapter 1

Introduction

"Other than skin cancer, breast cancer is the most common type of cancer among women in the United States. In 2003, an estimated 211,000 women will be diagnosed with breast cancer" (Natl. Inst. Health). Early detection is a major factor of survival. The most effective method of screening for breast cancer is mammography, which can detect a malignancy up to two years before a lump can be felt.

A digital mammogram generally detects to varying degrees the following signals/signs of breast cancer: clustered microcalcifications, speculated lesions, circumscribed masses, ill-defined masses, and architectural distortions. Many methods of analyzing digital mammograms have been recently examined and yielded varied success. Common techniques from the field of image processing have been applied to digital mammograms in an effort to locate signs of cancer sooner and more precisely than previously possible. Research suggests that computerized techniques applied/utilized by radiologists will be highly successful in analyzing digital mammograms. Computerized systems that draw attention to areas of suspicion, otherwise less noticeable to radiologists have the potential to greatly increase early detection.

1.1 Objectives of the Dissertation

The objective of this dissertation research is to address the needs of screening algorithms that are designed to aid in locating abnormalities in digital mammograms. Specifically, this dissertation seeks to provide and improve tools that are considered essential to the construction of a comprehensive screening system. This includes a mammogram

segmentation algorithm, a method for locating masses in a mammogram image, and a method for testing and comparing results from screening algorithms. In addition to functioning properly, the methods must be computationally efficient, i.e., their performance is feasible for use in a clinical environment.

Previously, an algorithm that effectively segments mammogram images into major sub-components and also meets the goals of efficiency and generality has been lacking. This research accomplishes this by creating an algorithm that segments a general mammogram image into its distinct components. This algorithm is computationally efficient and works on a general set of mammogram images without the need of "training sets." This new method is needed because it eliminates certain areas from consideration by other parts of screening algorithms. Also, effective image segmentation is required for focused and adaptive computer analysis of separate image components.

There are several techniques for locating masses in mammogram images. They function with varying levels of success, and many of them utilize a form of template matching to compare parts of an image with a selected model of a suspicious mass. Others construct feature vectors and employ a pre-trained classifier to choose malignant areas. This research seeks to bridge the differences between these types of methods and offer improvements in performance and accuracy. We also evaluate possible mass templates and choose an optimal template and a comparison method.

A significant shortcoming of many screening algorithms is that they are tested on limited numbers of images. This is usually due to the amount of time needed to evaluate, record, and compare the results of the algorithms. Much of this work must be done manually, including comparing computer results to pathology. We propose to create an

automated method for collecting, recording, comparing and reporting the results of mass location algorithms.

1.2 Contributions of the Dissertation

In this dissertation, we investigate a number of methods which are vital components of any mammogram screening algorithm. We have made improvements and enhancements to existing techniques and in some cases original techniques are developed. The specific contributions of this dissertation are summarized as follows:

- A mammogram image segmentation algorithm that effectively and efficiently divides mammogram images into distinct components.
- A mass location algorithm that identifies and extracts suspicious areas in mammogram images.
- A thorough examination and selection of an optimal template for detecting masses in mammogram images.
- A system for testing mass detection algorithms that allows automated comparisons to pathology.

1.3 Overview of Dissertation

Chapters two through four provide background in the areas of breast cancer, mammography, image processing, and computer assisted screening mammography. These are included to present prior research and also to give sufficient background for the original research presented in this dissertation. The strengths and weaknesses of current computer aided screening systems are also presented.

Chapter five presents a new and unique algorithm for segmenting mammogram images into the components of: the image background, breast, pectoral muscle, and subcutaneous fat. We provide detailed explanation of the algorithm with examples and results. This method is proposed as a pre-processing step in the screening process that allows for a significant improvement in algorithm efficiency by eliminating a large amount of image that must examined. This algorithm offers improved abnormality detection performance by reducing the potential for erroneous detections. Finally, the segmentation method presents the potential for use by adaptive detection algorithms by allowing the different components to be examined independently by further algorithms.

Chapter six describes the development of a method for locating areas in a mammogram image which contain masses. This technique uses a form of template matching to initially locate a suspicious area, and then applies a multi-scale step to further refine the detection(s). The algorithm for template matching is detailed and an evaluation of potential templates is presented. This method is also suggested as a pre-processing step that can extract suspicious areas for further analysis by a radiologist or other computer algorithm. It can quickly focus on limited areas in an image and reduce the search time needed for more complex analysis.

Chapter seven presents an algorithm for testing screening algorithms, recording the results of screening algorithms, and comparing those results with the pathology. An automated method for recording these screening results and comparing them to pathology is also included.

Chapter eight presents an experimental methodology for testing the algorithms presented. It also includes results from testing the techniques on mammogram images.

Finally, chapter nine presents a summary of the dissertation research. Significant contributions are highlighted, and additional avenues for research are outlined. Chapter nine also discusses the requirements and construction of a comprehensive computer aided screening system.

Chapter 2

Mammography

This chapter provides an introduction to breast cancer and a description of the screening methods and procedures of mammography. Examples of mammograms are presented with details. In addition, mammogram abnormalities are shown. This discussion is intended to provide sufficient background information and to demonstrate the need for computer aided screening algorithms in breast cancer detection.

2.1 Breast Cancer

> Cancer is a group of many related diseases that begin in cells, the body's basic unit of life…the body is made up of many types of cells. Normally, cells grow and divide to produce more cells only when the body needs them. This orderly process helps keep the body healthy. Sometimes, however, cells keep dividing when new cells are not needed. These extra cells form a mass of tissue, called a growth or tumor. Tumors can be benign or malignant. (Natl. Cancer Inst.)

Cells from benign tumors do not spread to other parts of the body and can be removed if necessary, although benign breast tumors are not life-threatening. Malignant tumors can invade and damage nearby tissues and organs, and spread to other parts of the body, a process called metastasis (Natl. Cancer Inst.).

Breast cancer refers to life-threatening malignancies that develop in one or both breasts and is the most common form of cancer among women in developed countries. In fact, one in eight women will develop breast cancer. Specifically in the United States, “the risk of breast cancer within five years among 50-year-old women was 0.8% (1 in 133) for Hispanic women, 0.9% (one in 107) for Asian/Pacific Island women, 1% (one in 98) for African American women and 1.3% (one in 75) among Caucasian women. Risks

increased with age, with the highest rates occurring in Caucasian women (8%, or one in 13) and African-American women (5.5%, or one in 18) at age 60" (Amer. Cancer Soc.).

The specific causes of breast cancer are not known. However, there are some risk-factors that may affect one's chances of developing breast cancer. They include early or late onset of menopause, family history of breast cancer, hormone replacement therapy, and certain dietary factors. However, in over 75% of women with breast cancer none of these risk factors are present (Hutt). In fact, the only clearly identifiable risks of developing breast cancer are gender and aging (Hutt). Given this lack of predictability, preventative measures have been very difficult to determine.

Breast cancer is either noninvasive, confined to the site of origin, or invasive, spreading. Noninvasive breast cancers include ductal carcinoma in situ (also called intraductal carcinoma or DCIS) and lobular carcinoma in situ, or LCIS. DCIS consists of cancer cells in the lining of the duct. DCIS is a non-invasive, early stage of cancer, but if left untreated, it may sometimes progress to an invasive, infiltrating ductal breast cancer. Although noninvasive, lobular carcinoma in situ is indicative of an increased risk of invasive cancer in both breasts (Univ. of Cal. Davis Heath System).

Invasive cancer occurs when cancer cells spread beyond the basement membrane, which covers the underlying connective tissue in the breast. This tissue is rich in blood vessels and lymphatic channels that are capable of carrying cancer cells beyond the breast. Invasive breast cancers include infiltrating ductal carcinoma and infiltrating lobular carcinoma. Infiltrating ductal carcinoma is an invasive breast cancer that penetrates the wall of a duct. It comprises between 70% and 80% of all breast cancer cases. Infiltrating lobular carcinoma is an invasive cancer that has spread through the wall of a lobule. It accounts for between 10% and 15% of all breast cancers. It may sometimes

appear in both breasts and in several separate locations (Univ. of Cal. Davis Heath System).

For treatment purposes, breast cancers have been divided into five stages. Stage zero describes cancer cells that are non-invasive but pose a long-term risk of becoming invasive. Stage One (I) describes tumors no more than 2 centimeters across that have not spread beyond the breast. In stage Two (II), the tumor is about 2 centimeters but has spread to the lymph nodes under the arm, or the tumor is about 5 centimeters but has not spread to the lymph nodes under the arm. Stage Three (III) cancers are more than 5 centimeters across and have spread to lymph nodes or other tissue near the breast. Stage Four (IV) cancers are called metastasized cancers and have spread to other parts of the body. The seriousness of the diseases increases with each stage as the survival rates decrease.

Many consider early detection the key to surviving breast cancer. The most effective method of screening is mammography, which can detect a tumor up to two years before a lump can be felt. A mammogram generally detects to varying degrees the following signals/signs of breast cancer: clustered microcalcifications, spiculated lesions, circumscribed masses, ill-defined masses, and architectural distortions.

In recent years the occurrence of breast cancer has risen, but the death rate has fallen. This is largely attributed to the ubiquitousness of mass screening programs.

> Between 1950 and the late 1980s, overall death rates from breast cancer were relatively stable, according to the ACS publication, Breast Cancer Facts & Figures 2001-2002. The death rates for breast cancer then began to fall, dropping by about 1.6% each year between 1989 and 1995. Between 1995 and 1998, the drop in the rates picked up, declining about 3.4% each year. (Amer. Cancer Soc.)

A recent study looked at breast cancer death rates among Swedish women during a time span in which the government made mammograms widely available (Cady and

Michaelson). Among women who had periodic screening mammograms, deaths from breast cancer dropped by 63%, compared to 10 years before, when widespread mammography was not available. Death rates did not drop significantly for women who did not have periodic screening mammograms. While new treatments and greater awareness of palpable lumps have improved a woman's chances for surviving breast cancer, screening mammography is considered an essential tool for early detection.

2.2 Screening Mammography

By definition, screening mammography is performed on patients with no signs of breast cancer to help detect cancer early. Early detection is a key factor in successful treatment of breast cancer.

2.2.1 The Mammography Process

Mammography is an X-ray technique for examining the breast. Like other radiographic examinations, an X-ray beam is passed through the tissue to record the variations in amounts of radiation that are absorbed. Since different tissues in the breast absorb different amounts of radiation, it is possible to distinguish features and details about the tissues being examined. In screening mammography each breast is compressed into a relatively flat surface. Then an X-ray source on one side of the breast emits radiation through the breast. On the other side of the breast the radiation is recorded on film or by an electronic device. In a typical screening mammogram, each breast is examined from an overhead view, called craniocaudal, and from the side, called mediolateral. Figure 2.1 shows an example of a craniocaudal view from a mammogram. Figure 2.2 is the mediolateral view from the same mammogram and of the same breast as shown in Figure

2.1. The standard practice of taking two views of each breast has been shown to be more sensitive and effective in detecting signs of cancer (Hutt).

2.2.2 Mammographic View of Breast Anatomy

The interior of the female breast contains mostly fatty and fibrous connective tissues. It is divided into about twenty sections called lobes, which are further subdivided into lobules. These are structures that contain milk-producing glands. These glands secrete milk into a system of ducts, which carry the milk through the breast and converge in a collecting chamber located just below the nipple. Figure 2.3 shows a mediolateral mammogram image in which the basic components of the breast are labeled.

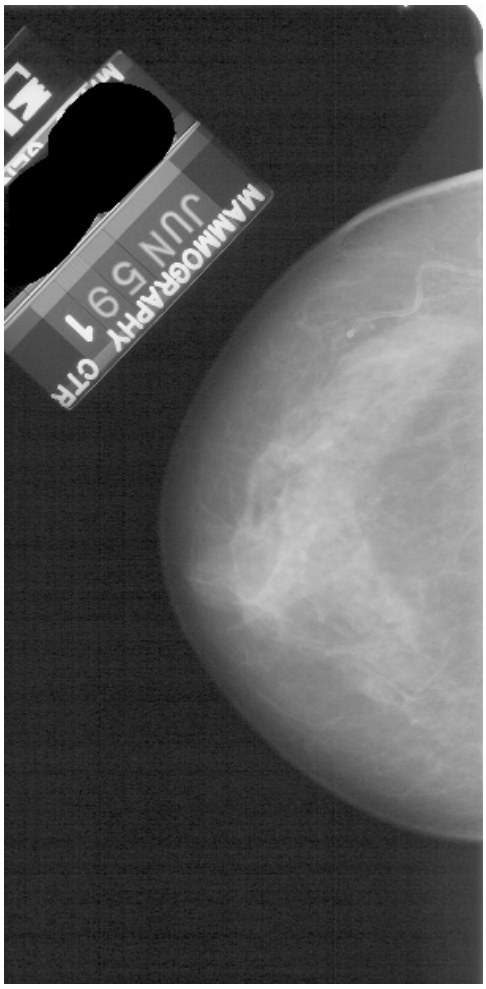


Figure 2.1: Craniocaudal view of left breast in mammogram image

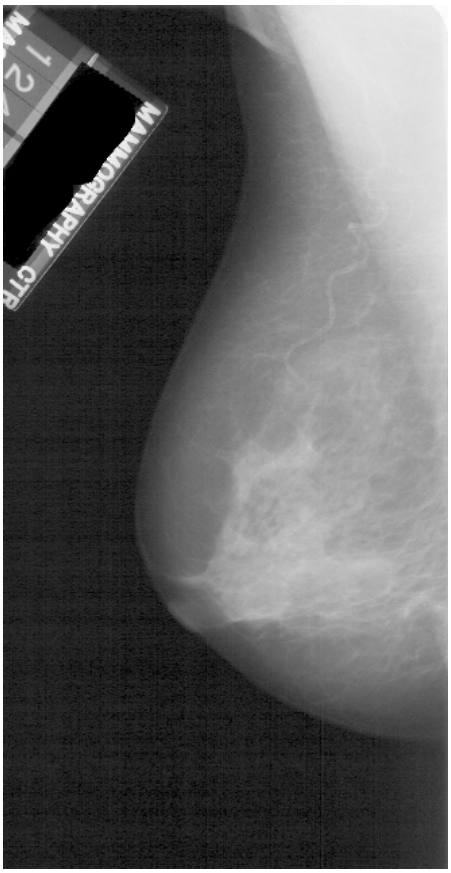


Figure 2.2: Mediolateral view of left breast in mammogram image

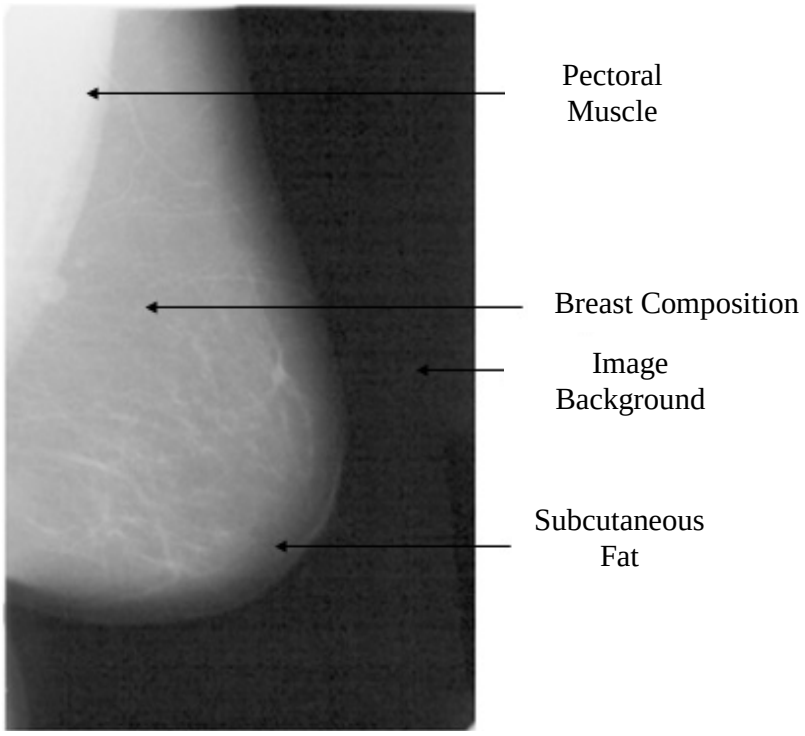


Figure 2.3: Labeled mammogram image

Denser tissues in the breast display brighter intensities in the mammogram images. Muscles, fibroglandular tissue, malignant and benign masses, and vascular tissue appear brighter. Areas containing fat or skin appear darker.

2.2.3 Objects in Mammogram Images

In addition to the basic anatomy of the breast, a variety of objects appear in mammogram images. These include masses, microcalcifications, and architectural distortions.

A mass is a space-occupying lesion. If seen in only one projection it is referred to as a density. It is referred to as a mass only if viewed in both projections (D'Orsi). There are several types of masses found in mammogram images. Masses are categorized by their shape, density, and margins. The shapes include: round, oval, lobular, and irregular. The margins include: circumscribed, microbulated, obscured, indistinct and spiculated. The densities include: high density, low density, equal density, and fat containing. These categories help radiologists to precisely describe masses found in mammograms and to classify masses as benign or potentially malignant. The focus of this research is the location of masses in mammogram images and therefore masses will be discussed at length and pictorial examples presented. Figures 2.4-2.7 illustrate some mass shapes found in mammogram images. Round masses are generally circular and oval masses are elliptical in shape. Lobular masses display contours and undulations. Masses with shapes that cannot be characterized are termed irregular.

Figures 2.8-2.12 show the varying types of mass boundaries observed in mammogram images. Circumscribed masses display distinct well-defined boundaries. Microlobulated masses have margins, which undulate in small cycles. Obscured masses are those that are hidden by superimposed or adjacent normal tissues. Indistinct masses

have poorly defined boundaries that taper into the background. Spiculated masses have spoke-like lines radiating out from the mass.

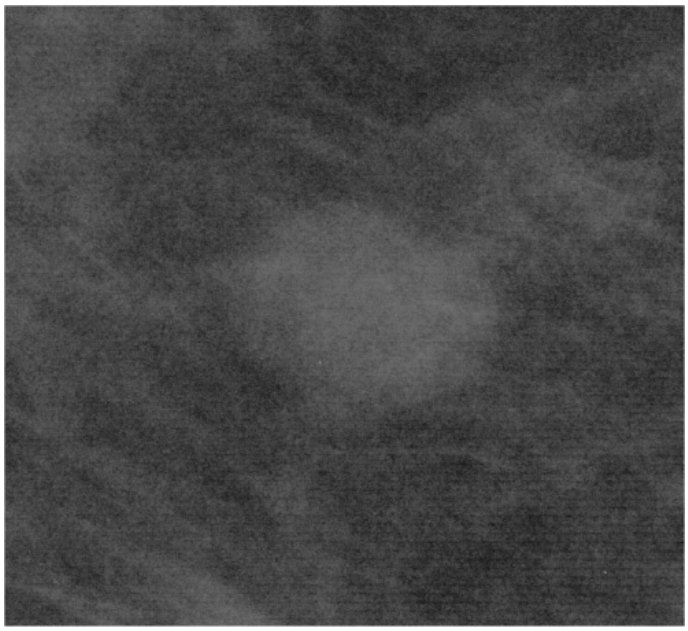


Figure 2.4: Example of round mass

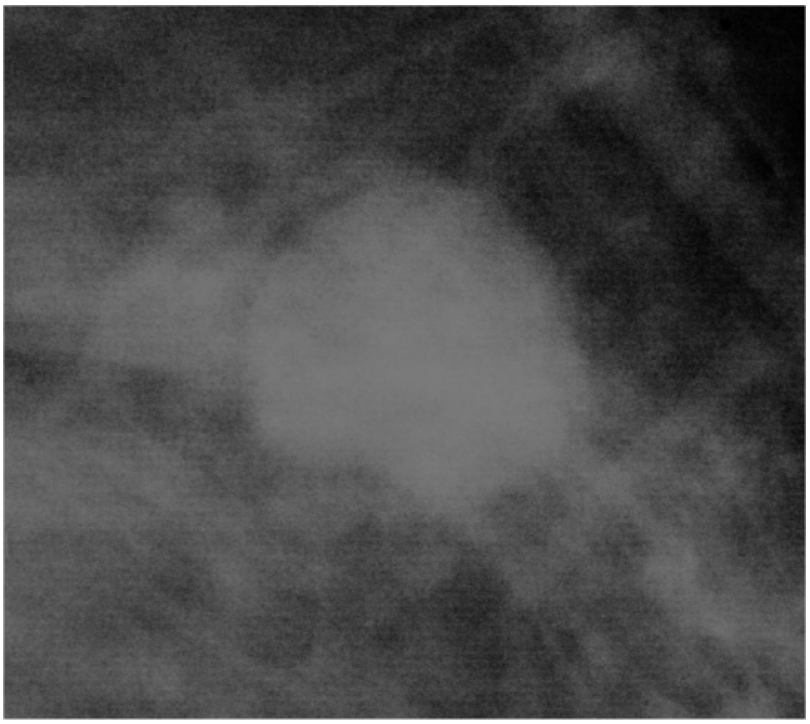


Figure 2.5: Example of oval mass

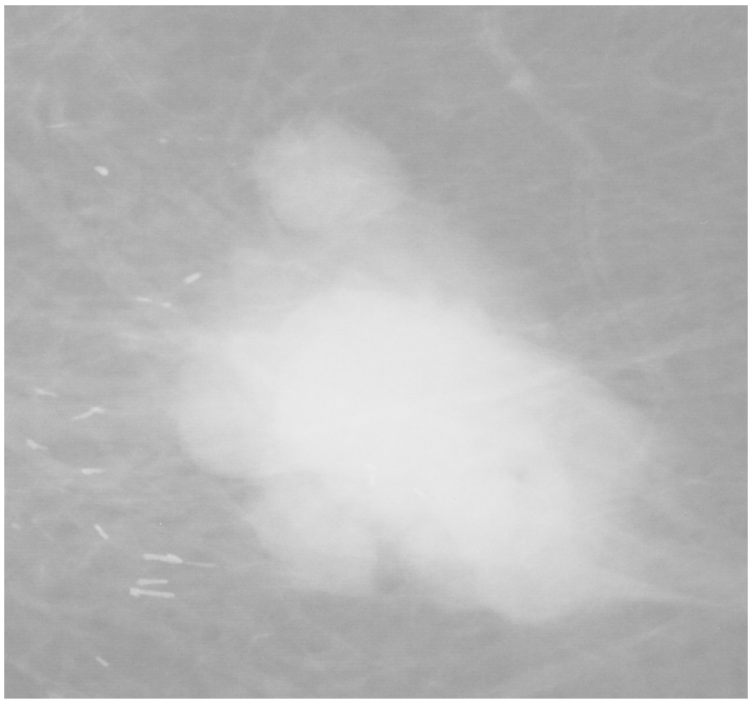


Figure 2.6: Example of lobular mass

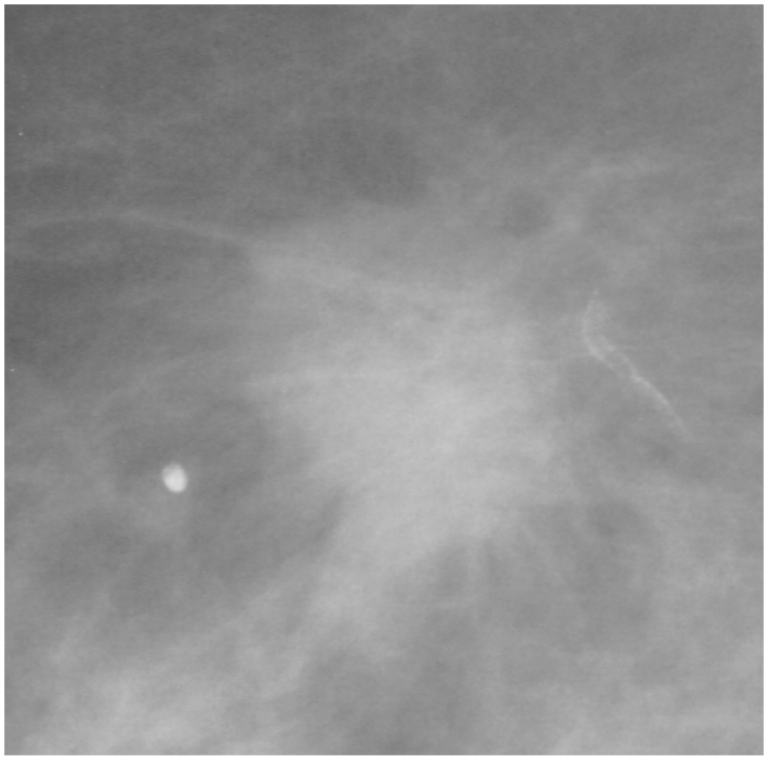


Figure 2.7: Example of irregular mass

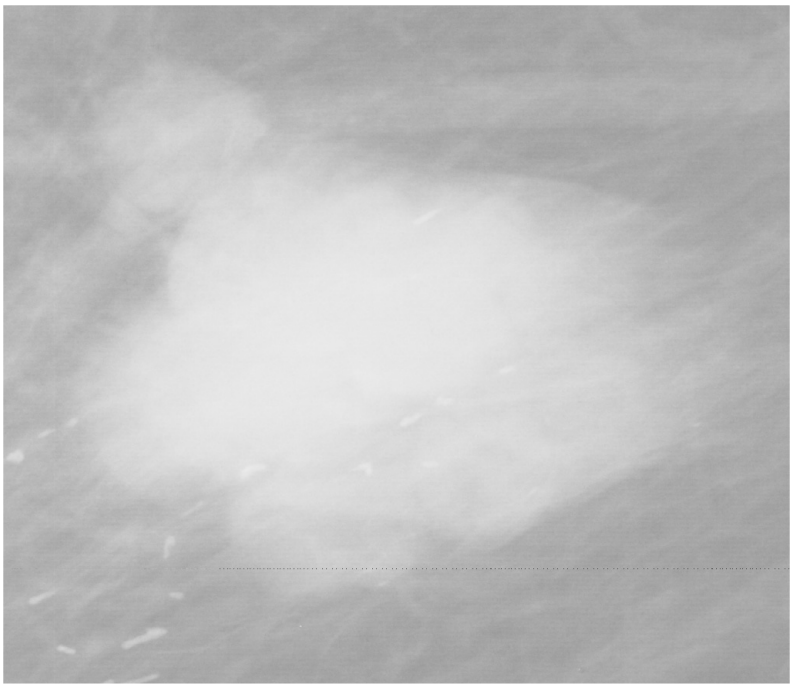


Figure 2.8: Example of circumscribed mass

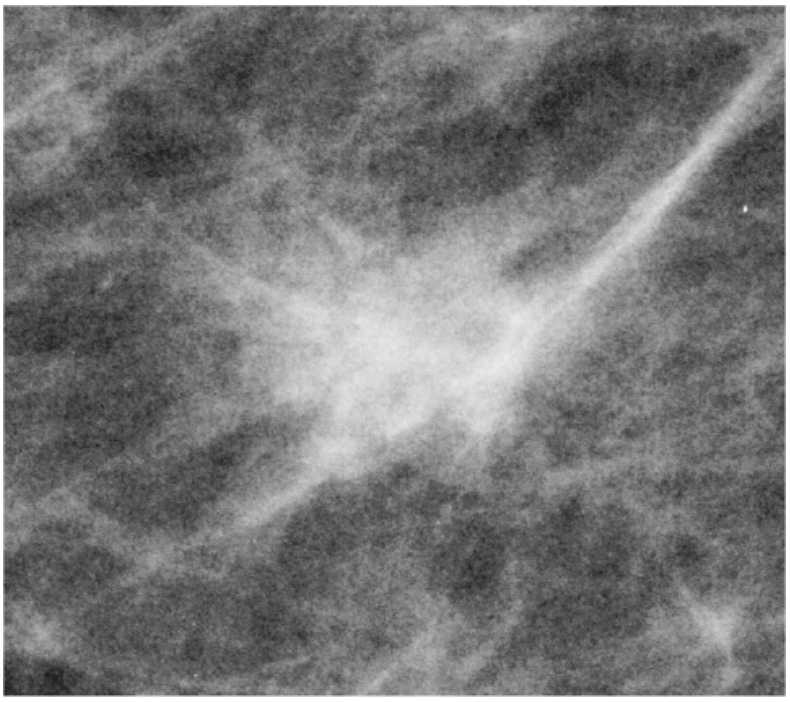


Figure 2.9: Mass with microlobulated boundaries

Figure 2.10: Mass with obscured boundaries

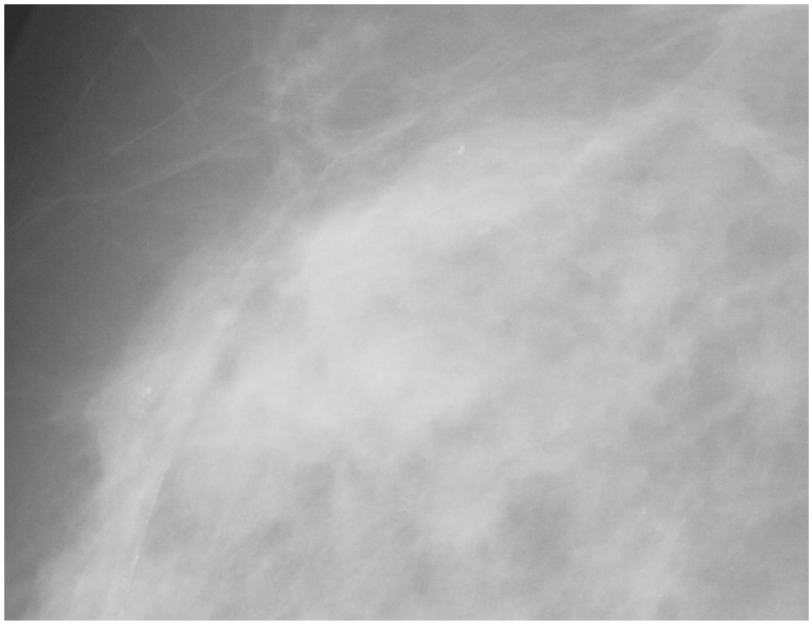


Figure 2.11: Mass with indistinct boundaries

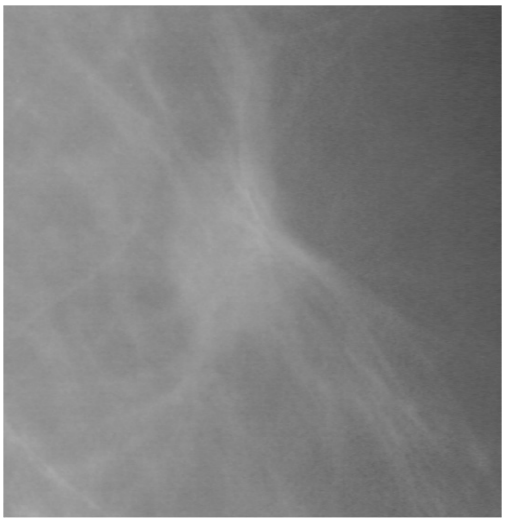


Figure 2.12: Example of spiculated mass

Figures 2.13-2.15 present examples of masses with varying densities. High density masses stand out sharply from the background tissue. Low density masses are faint and can be partially obscured. Equal density masses display similar intensities as other objects in the image. Fat containing masses are faint due to low-density fat deposits in the mass.

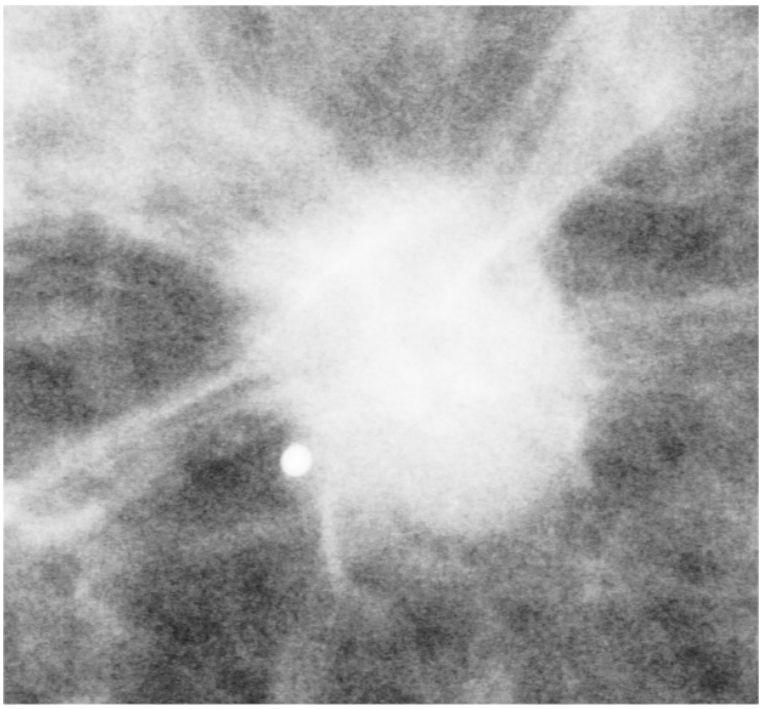


Figure 2.13: Example of high density mass

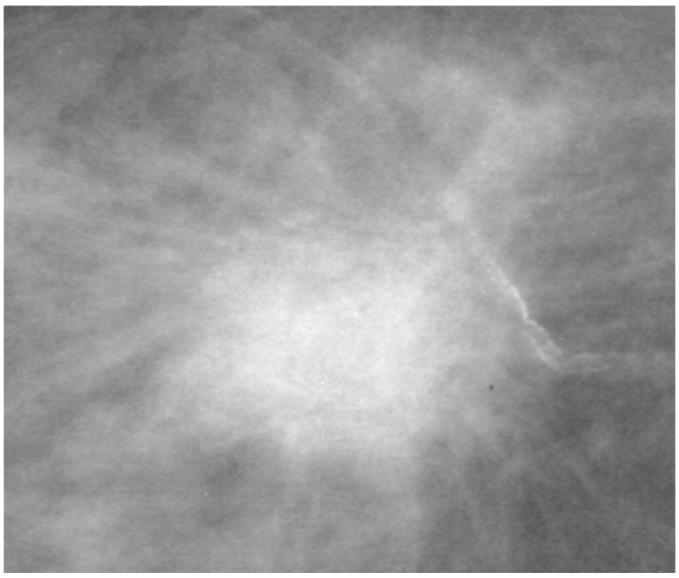


Figure 2.14: Equal density mass

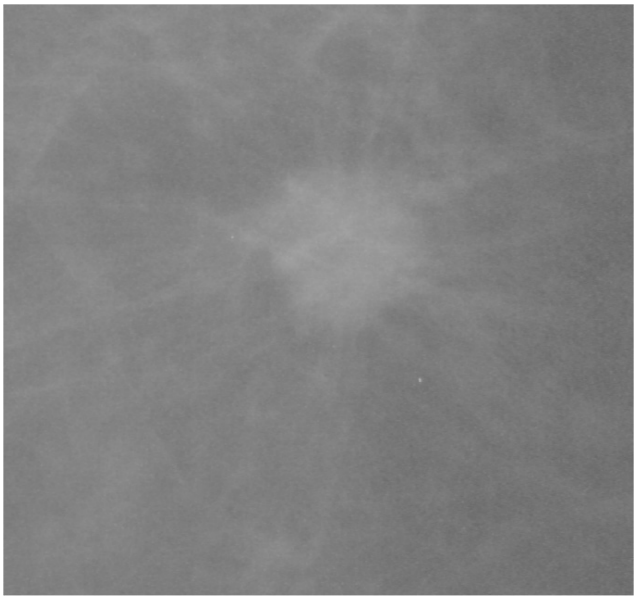


Figure 2.15: Low density mass

Other important objects observed in mammograms are microcalcifications. These represent calcium deposits located in the breast tissue and are considered highly indicative of breast cancer. Microcalcifications appear as small, bright objects that stand out from the surrounding tissue. Figure 2.16 shows several examples of microcalcifications.

A final category of abnormalities observed in mammogram images is architectural distortions. These are instances in which the normal structure of the breast is observably distorted. This includes spiculated areas and/or retraction from a focal point. These distortions are seen even though no mass is visible. Figure 2.17 shows examples of architectural distortions.

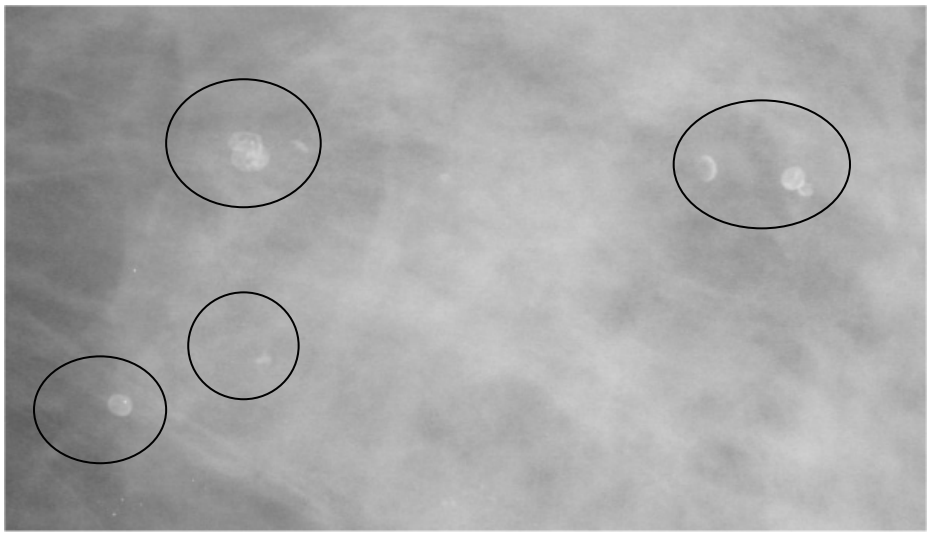


Figure 2.16: Microcalcifications (circled)

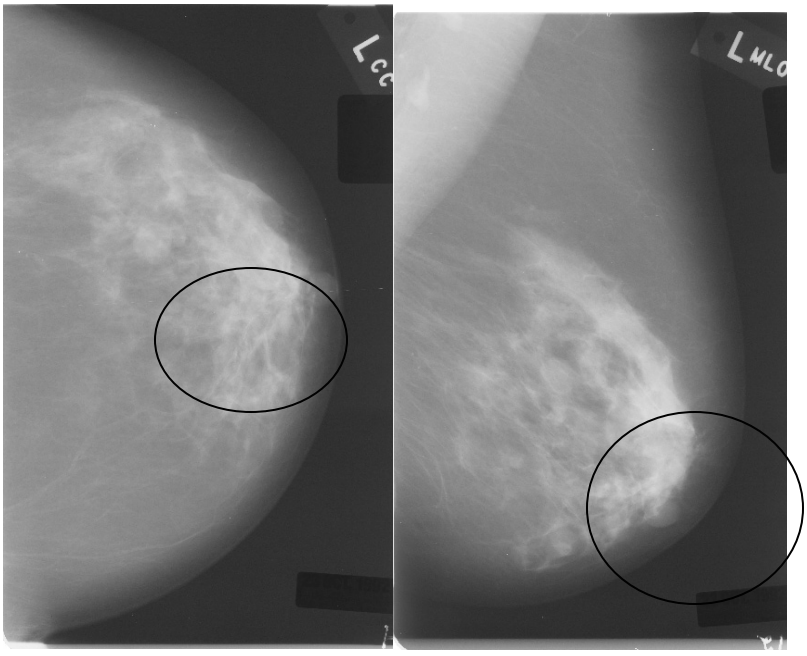


Figure 2.17: Side and overhead view of an architectural distortion (circled)

2.2.4 Background Observations in Mammogram Images

In addition to specific abnormalities seen in mammogram images, other phenomena are observed which may affect the examination of the mammogram by the radiologist. For example, some mammogram images show breasts which are very dense. They display dense fibroglandular tissue, which can make it difficult to detect cancerous signs. Figure 2.18 shows a mammogram image which displays dense tissue.

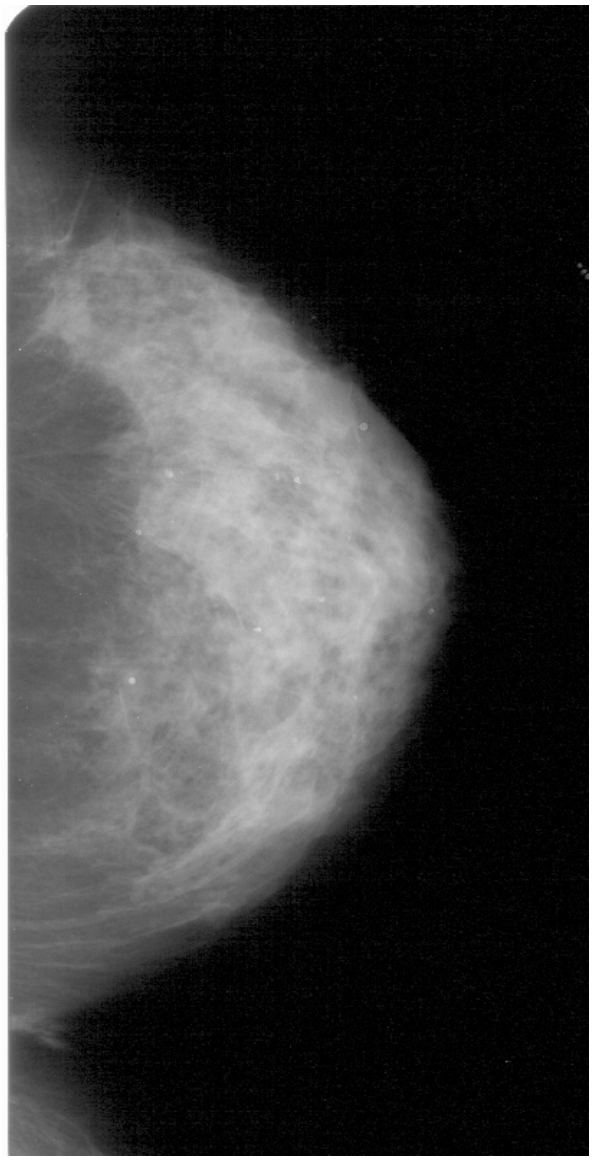


Figure 2.18: Mammogram with very dense tissue

2.3 Mammogram Film Reading

2.3.1 Reading Errors

Screening mammography can help early detection of breast cancer; however, this is dependant upon proper interpretation of the mammogram by a radiologist. Because of the subtleties and variations of the breast, errors can be common. There are two errors typical in examining mammograms. They are false positives and false negatives. False positives are instances where the radiologist identifies an area of the breast as cancerous when it is benign. False-negatives occur when an abnormality is not detected by the radiologist.

False positives are the less severe of the two errors. They typically do not endanger the life of the patient, but they do have negative consequences. Additional mammogram and/or more invasive tests are required to determine the nature of the abnormality.

False-negatives can be very serious. They directly delay or prevent early detection and can adversely affect the woman's chances of surviving breast cancer. Tumors or signs of tumors that are not detected or misclassified as benign reduce the effectiveness of screening mammography.

Some researchers have classified radiological examination errors into three categories: search errors, recognition errors and decision-making errors (Hutt). Search errors refer to instances in which the examining radiologists simply did not examine the area containing a significant abnormality. Recognition errors occur when the areas containing abnormalities are searched, but the abnormalities are not identified. Decision-making errors are those in which an abnormality was located but misclassified as benign.

2.3.2 Prompting Systems

Research has shown that successful location of abnormalities depends on the location of the abnormality falling into the direct field of view of the examining radiologist (Gale and Worthington). Furthermore, Kundel and Nodine have demonstrated that in a number of cases, abnormalities were overlooked because they were not specifically attended to. This implies that if radiologists were directed to specific areas that contained abnormalities, they would more readily detect abnormalities.

In a study conducted by Chan and colleagues, an automated abnormality detection system was used to produce prompts for the radiologist examining a mammogram. This type of system is called a prompting system. Prompting systems are designed to aid a radiologist by highlighting areas of suspicious content, those that may contain an abnormality. While no computer algorithm can be used to diagnose breast cancer or specific signs of breast cancer, they can be used to assist the radiologists in examining mammograms. Chan and others showed that the use of a prompting system, which presented radiologists with areas of suspicion, yielded a significant increase in detection performance of abnormalities by examining radiologists.

Chapter 3

Image Fundamentals

This chapter introduces techniques and tools in digital image processing which have been used to analyze and diagnose digital mammograms. Several algorithms and principles are presented with examples applied to digital mammograms. These techniques, operations and statistics form the tools for development of comprehensive analysis/diagnosis algorithms.

3.1 Digital Mammograms as Images

Most mammograms are currently recorded on film. However, some practices now use equipment with digital acquisition capabilities. The images used in this research are from mammogram films which were digitized by a scanner.

3.1.1 Digital Images

> A digital image is an image $x_{ij}=f(i,j)$ which has been discretized both in spatial coordinates and in brightness. We may consider a digital image as a matrix whose row and column indices identify a point in the image whose corresponding matrix elemental value identifies the gray level at that point. The elements of such a digital array are called image elements, picture elements, pixels or pels. (Gonzales and Wintz)

This is the definition of a gray level image, also called a grayscale image. The images used in this research are grayscale images. Grayscale images can be used to show variances in relative intensity for a given scene or subject matter. Because the intensities captured on a mammogram x-ray film are records of the relative absorption of radiation, grayscale images are entirely suitable for digital mammogram images. The elements in a

digital image contain a discrete value, usually a positive integer within a given range. Typically images will be defined by the range of values they contain. For example, an eight-bit grayscale image is one in which the pixel values range from 0 to 255. A twelve-bit grayscale image contains pixel values ranging from 0 to 4095. Likewise, a binary, or one-bit, image contains pixels which have values of zero or one.

3.2 Image Statistics

Numerous statistics can be developed from digital images which aid in describing and analyzing images. This section provides an introduction to a selected group of image statistics with examples. The selected group represents those used in this research; however, they are far from exhaustive. In fact, a significant amount of research continues developing new statistics for describing and analyzing images. This research is primarily interested in the application of existing statistics.

The following statistics can be computed over an entire image or over a sub-image. In many cases, multiple sub-images from a single image are selected and compared using these statistics.

- **Histogram**

The histogram of an image represents the distribution of gray level intensities in an image irrespective of location in the image. Figure 3.1 is an example mammogram image and Figure 3.2 is its corresponding histogram. The histogram shows the probability distribution for each gray level in the image. Since Figure 3.1 is an eight-bit image, the histogram shows 256 values.

Figure 3.1: Example mammogram image representing side view of left breast

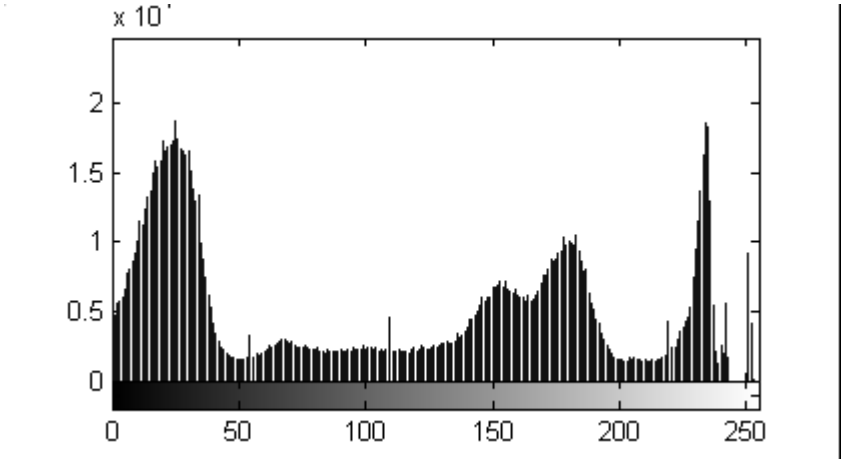
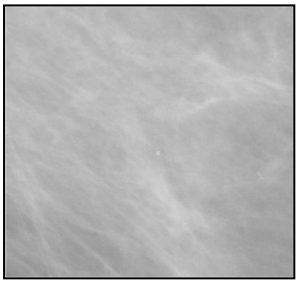


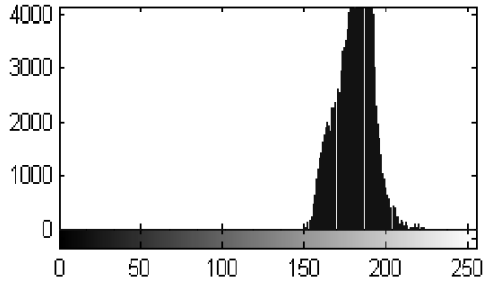
Figure 3.2: Histogram of image in Figure 3.1

The histogram of an image is computed by recording the number of times each pixel's intensity within the valid range appears in the image. For an eight-bit image, the histogram will contain 256 values.

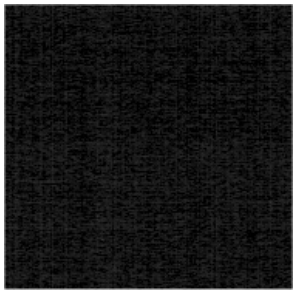
Much information can be gathered from the histogram of an image and sub-images. For example, Figure 3.3 shows three sub-images from Figure 3.1 and their corresponding histograms. The three sub-images represent adipose breast tissue, Figure 3.3(a), image background, Figure 3.3(c), and pectoral tissue, Figure 3.3(e). The histograms of the sub-images provide a quantitative view of the obvious visual differences between the mammogram images components represented.



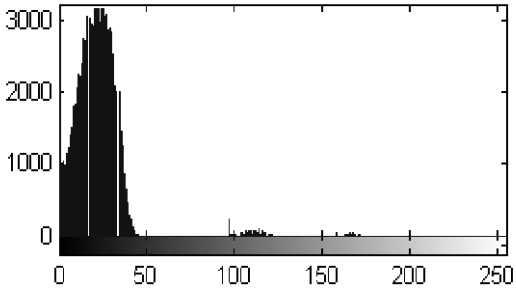
(a) Adipose breast tissue



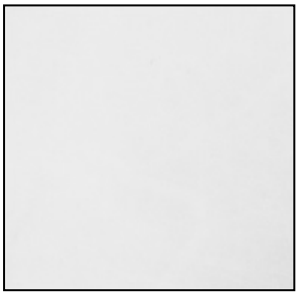
(b) Histogram of (a)



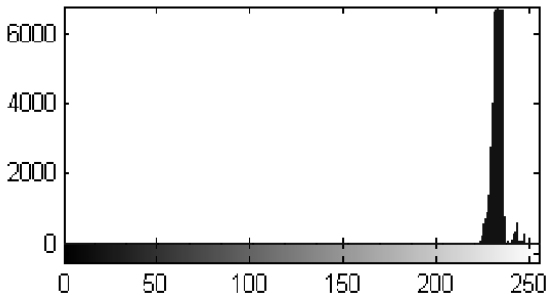
(c) Image background



(d) Histogram of (c)



(e) Pectoral tissue



(f) Histogram of (e)

Figure 3.3: Selected sub-images and corresponding histograms

Note that the histogram of a discrete image can serve as the probability mass function (pmf). The pmf is expressed as p(x), where x is not a spatial coordinate, but is an intensity within the valid range for the image. The value $p(x_{ij})$ is equal to the number of times that the intensity x_{ij} appears in the image divided by the total number of pixels in the image. The value x_{ij} varies from the minimum to the maximum intensity value. This means that $p(x_{ij})$ equals the probability that a given pixel will be of value $x_{ij.}$ At this point, it should be noted that the pdf and its accompanying statistics exist irrespective of the spatial location of the individual pixels.

- **Mean**

The mean of an image is the average value of all the pixel intensities in the image. This can be computed over the whole image or over a sub-image. The mean is defined by:

$$\mu \equiv \frac{\sum_{i,j} x_{ij}}{N} \tag{3.3}$$

N is the total number of pixels in the image. The mean is sometimes also called the expected value. Expected value is defined as:

$$E[x] \equiv \sum_{i,j} x_{ij}\, p(x_{ij}) \tag{3.4}$$

Also, the expected value can be applied to functions of x_{ij}. The median of an image is related to the mean and is defined as the middle value of all the pixels. So, if all the pixel values in the image were sorted, the middle value would be the median. The mode of an image is the pixel value which occurs with the greatest frequency.

- **Variance/Standard Deviation**

The variance of an image is a measure of the variation of pixel intensities in the image. The variance of an image, $x_{i,j}$, is defined:

$$\sigma^2 \equiv \frac{\sum_{i,j} \left(x_{i,j} - \mu\right)^2}{N} \tag{3.5}$$

where μ is the mean of the image and N is the number of pixels. σ^2 is the variance and σ equals the standard deviation. Variance can also be defined in terms of the expected value of x_{ij} as follows:

$$\sigma^2 \equiv E[(X - \mu)^2] \tag{3.6}$$

- **Skewness**

Skewness is a measure of the asymmetry of the pixel values around the image mean. If skewness is negative, the pixel intensities are spread out more to the left of the mean than to the right. If skewness is positive, the pixel intensities are spread out more to the right of the mean. The skewness of a distribution is defined as:

$$y \equiv \frac{\sum (x_{i,j} - \mu)^3}{N\sigma^3} \tag{3.7}$$

where μ is the mean of the image, σ is the standard deviation, and N is the total number of pixels. The skewness of any perfectly symmetric distribution is zero.

- **Moments**

Moments are special cases of the expected value and can be useful descriptions of the image and its intensity distribution. They are used as indications of the shape of the intensity distribution of the image. The *r*th moment is defined as:

$$\mu_r' \equiv \sum_{i,j} x_{ij}^{\ r}\, p(x_{ij}) \tag{3.8}$$

The first order moment, μ_1', is the mean. In many cases the moments are centered around the mean. This is called the central moment and is denoted as follows:

$$\mu_r \equiv \sum_{i,j} (x_{ij} - \mu)^r\, p(x_{ij}) \tag{3.9}$$

The 2nd central moment, by definition, is the variance.

- **Kurtosis**

Kurtosis is a measure of whether an image's intensity distribution is peaked or flat relative to a normal distribution. That is, image distributions with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Image distributions with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case. The formula for kurtosis is:

$$kurtosis \equiv \frac{\sum_{i,j} (x_{i,j} - \mu)^4}{(N-1)\sigma^4} \tag{3.10}$$

The kurtosis for a standard normal distribution is three. Positive kurtosis indicates a "peaked" distribution and negative kurtosis indicates a "flat" distribution (Natl. Inst. of Statistics).

- **Center of gravity**

The center of gravity of an image or sub-image provides a method for attaching spatial significance to the mean of an image. The center of gravity of an image is the location in the image that any line drawn through the image intercepting that point divides the image into two parts each having the same sum of intensities. It is computed as:

$$cog_i \equiv \frac{\sum_{ij}(x_{ij} * i)}{N}, cog_j \equiv \frac{\sum_{ij}(x_{ij} * j)}{N} \quad (3.11)$$

- **Correlation**

Statistical correlation provides a method for comparing the "likeness" of two datasets, signals or images. It is defined as follows:

$$cor(x, y) \equiv \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} \quad (3.12)$$

where $\mathrm{cov}(x, y) \equiv \mu_{xy} - \mu_x \mu_y$.

Figure 3.4 shows an example image which contains several geometric shapes. Figure 3.5 is a template image of a circle. Using statistical correlation between the template image and each pixel in the example image, one can create a new image which shows the locations of high and low correlation between the template and the example. Figure 3.6 shows the resulting correlation image. Areas in Figure 3.6 with high intensity correspond to circles in the example image. This is a useful method for locating specific features in images.

Correlation is often used in template matching. It can compare a template or prototype image with the image being examined and locate areas of similarity. A simplistic example is when one uses a photo of a train as a template and performs correlations over an aerial photograph; areas that contain trains would have higher correlations that those that did not.

These statistics allow concise, quantified descriptions of images, which are typically discriminated visually. They will be used extensively in this research to analyze and compare components of digital mammogram images.

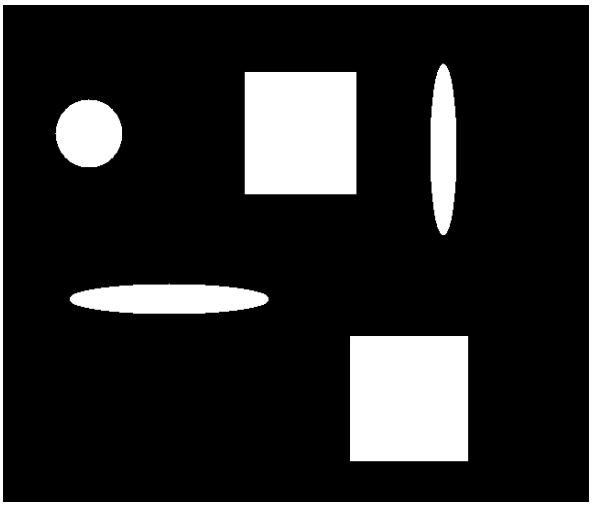


Figure 3.4: Example image with geometric shapes

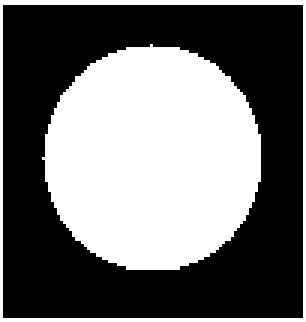


Figure 3.5: Circle Template

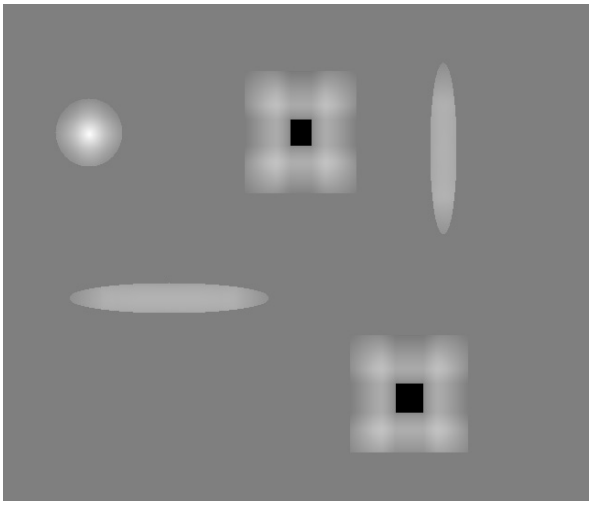


Figure 3.6: Result from correlation

3.3 Image Enhancement

This section presents selected methods for enhancing digital images. Image enhancement operations can be used to improve the appearance of images, to eliminate noise or error, or to accentuate certain features in an image. These operations can be quite useful in the development of digital mammogram analysis techniques.

3.3.1 Histogram Operations

Section 3.2 details the construction of the histogram. The histogram describes the data in the image, but it can also be used to enhance the image. Operations can be performed on the histogram to improve the appearance of the image or to adjust the image content.

- **Histogram Scaling**

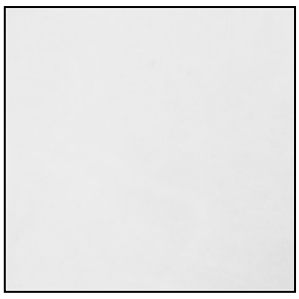
Some images display low contrast, i.e., they display little difference between their brightest and darkest pixels. In other words, they utilize a small proportion of the available intensity range. This frequently occurs when analyzing sub-images from digital mammogram images. Figure 3.7(a) shows a sub-image from the mammogram image in Figure 3.1. This sub-image presents an area of the pectoral muscle; note the low contrast. The histogram shown in Figure 3.7(b) is very narrow. Features in this sub-image are almost impossible to distinguish, and the pixels in this sub-image vary from 221 to 250. This represents 11.37% of the possible 255 values for this eight-bit image. Figure 3.7(c) shows the image after the pixel values have been scaled to the entire intensity range of 0 to 255. Figure 3.7(d) shows the resulting histogram. This image is much broader and has more detail and features.

This is an example of simple linear scaling of pixel intensities, in some cases more complex operations provide better results. The intensities can be scaled to add

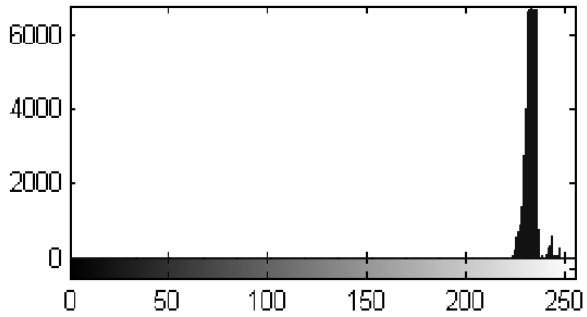
contrast to various bands of the intensity spectrum. Scaling can focus on image features which fall in a certain intensity range.

- **Histogram Equalization**

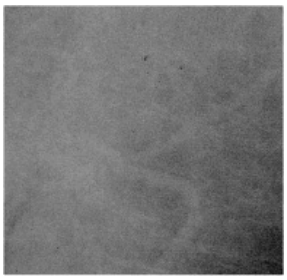
Another method to enhance the contrast of an image is called histogram equalization. Histogram equalization is a process used to produce a new enhanced image with uniform histogram. This is achieved by using a normalized cumulative histogram as a gray scale mapping function. Figure 3.8(a) shows a sub-image from Figure 3.1 displaying a portion of the pectoral muscle. As described earlier, it has poor contrast. Figure 3.8(c) shows the same sub-image after historgram equalization. Figure 3.8(d) shows the resulting histogram, and the details and features in the sub-image are more visible.



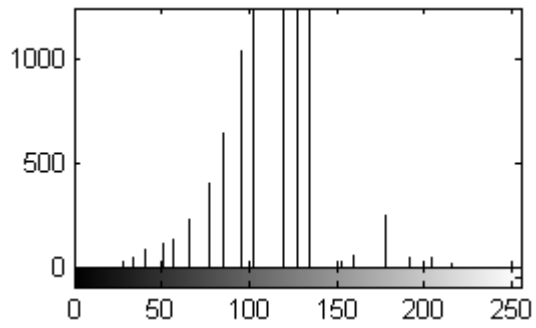
(a) Pectoral tissue



(b) Histogram of (a)

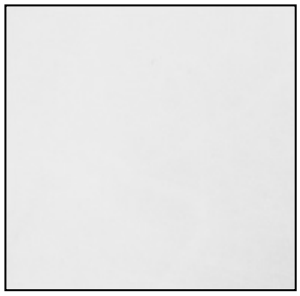


(c) Sub-image in (a) with scaled histogram

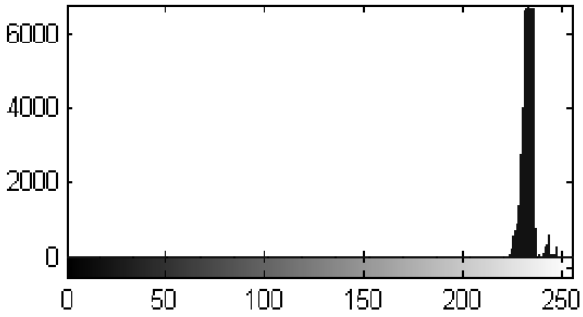


(d) Histogram of (c)

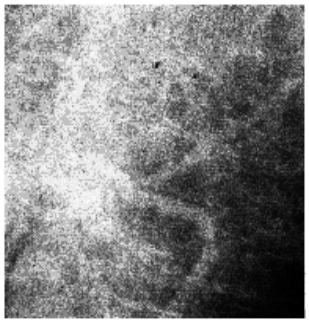
Figure 3.7: Example of histogram scaling for contrast enhancement



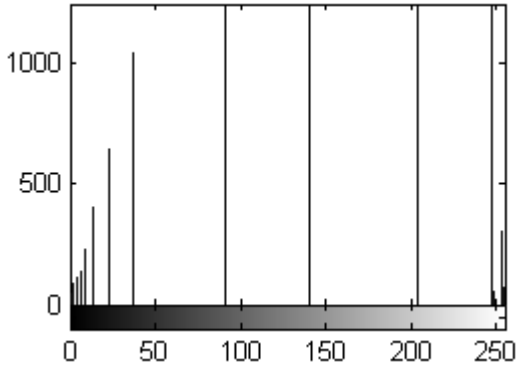
(a) Pectoral tissue



(b) Histogram of (a)



(c) Sub-image in (a) with equalized histogram



(d) Histogram of (c)

Figure 3.8: Example of histogram equalization

3.3.2 Filtering and Convolution

Linear filtering an image can alter an image in a variety of ways, to enhance or suppress certain features. The difference between linear filtering and other methods of image enhancement is the use of a filter, also called a kernel. The linear filter is accomplished through an operation called convolution. In convolution, the value of an output pixel is computed as a weighted sum of neighboring pixels. The matrix of weights is called the convolution kernel. For example, applying a convolution kernel K=[1/3 1/3 1/3] to a one dimensional signal A=[1 0 1 0 1 0 1 0 1 0] has the effect of replacing each value with the

average of the value, the value to left, and the value to right. This yields the result B=[1/3 2/3 1/3 2/3 1/3 2/3 1/3 2/3 1/3 2/3].

3.3.3 Image Smoothing

In some cases an image will display more detail than needed. If one is analyzing an image for large scale features that are composed of many pixels, the minor variations between pixels can pose a distraction. To reduce the small scale features in an image, we will use image smoothing. There are several methods for smoothing images. The most common is the use of an average filter. The filtering example in the previous section was an averaging filter. An averaging filter can be thought of as an integration across the image. Figure 3.9 shows a mammogram image with many fine details. Figure 3.10 shows the image in Figure 3.9 after 20x20 averaging filter has been applied to reduce the detail.

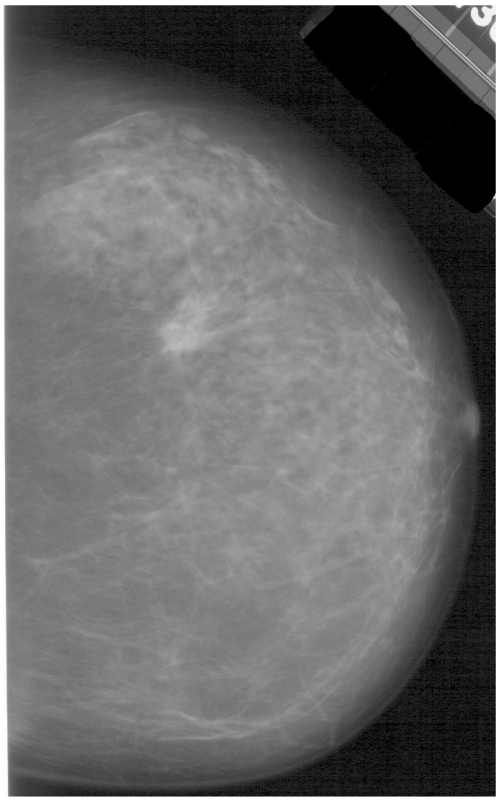


Figure 3.9: Mammogram image with fine details

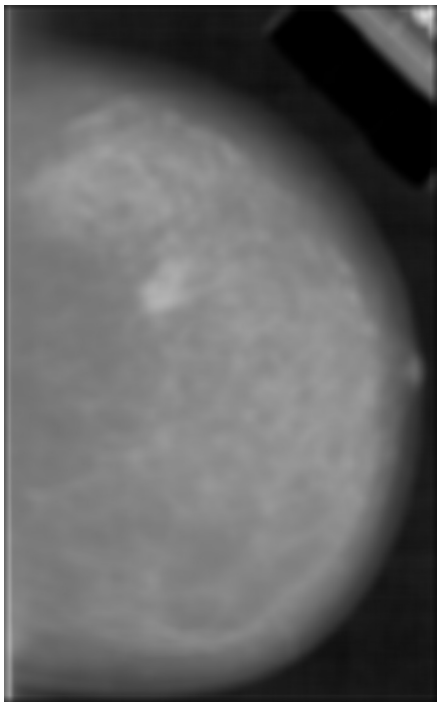


Figure 3.10: Image from Figure 3.9 after 20x20 average filter

3.3.4 Image Sharpening

The opposite of image smoothing is called image sharpening and can be thought of as image differentiation (Gonzales and Wintz). What is needed for this process is a kernel which will perform as a differential operator - an operator that enhances the differences between neighboring pixels. In many cases, the kernel chosen is an approximation of the Laplacian operator. The Laplacian operator is defined as:

$$\nabla^2 f(i,j) \equiv \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{3.13}$$

A useful approximation of Formula 3.13 is:

$$\nabla^2 f(i,j) \approx f(i+1,j) - 2f(i,j) + f(i-1,j) + f(i,j+1) - 2f(i,j) + f(i,j-1) \tag{3.14}$$

which can also be written as:

$$f(i+1,j) + f(i-1,j) + f(i,j+1) + f(i,j-1) - 4f(i,j) \tag{3.15}$$

This is equivalent to the convolution kernel:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.16}$$

The approximation can be expanded to include more terms, but for illustration the given kernel is sufficient. The Laplacian operator is useful for enhancing edges in an image. Figure 3.11 is the result of applying the Laplacian kernel 3.16 to the image on Figure 3.10; this presents the edges in the image. Figure 3.12 shows the results of those added edges back into the original image (Figure 3.9).

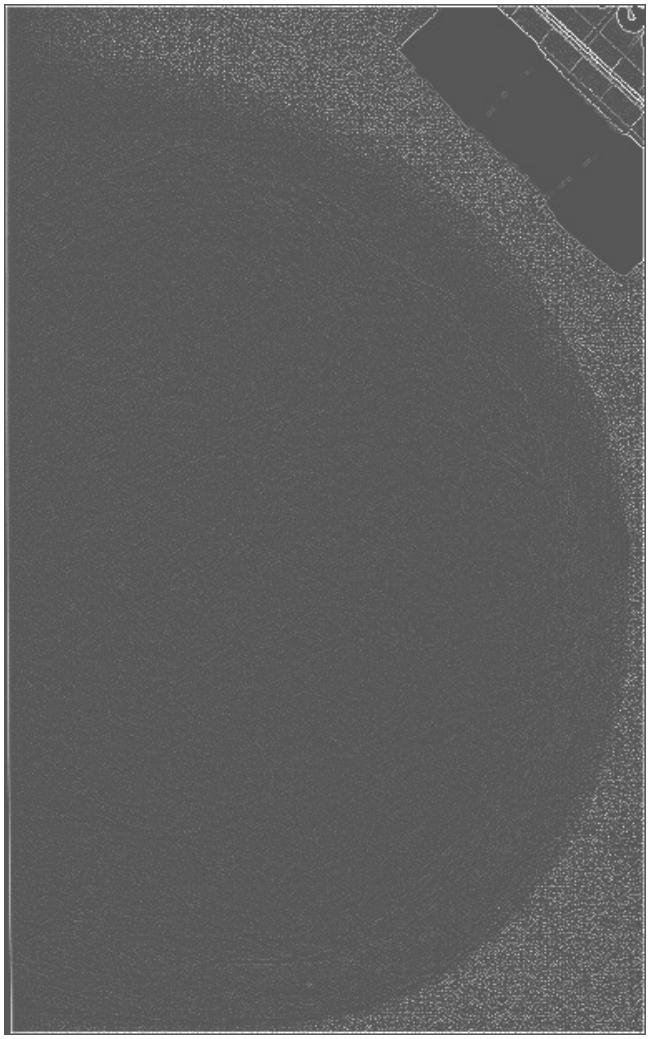


Figure 3.11: Laplacian detected edges in mammogram image

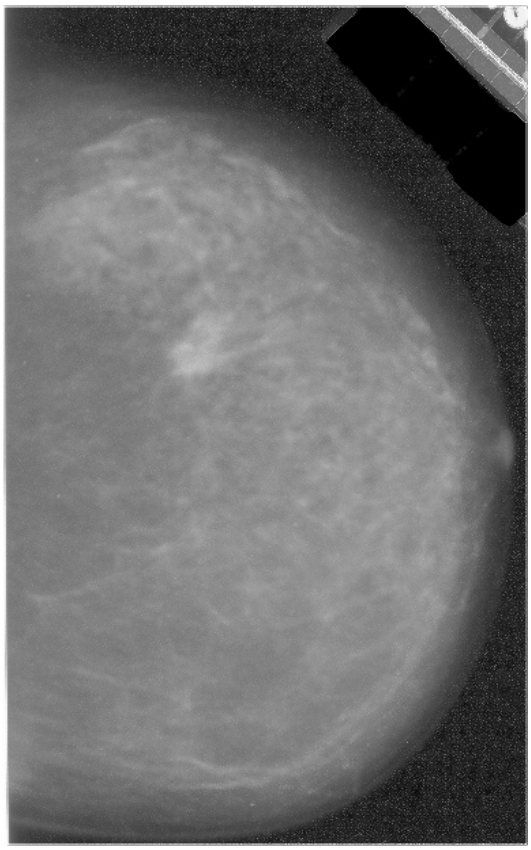


Figure 3.12: Mammogram image with edges enhanced

3.4 Image Segmentation

The process of dividing an image into distinct, meaningful regions is called image segmentation. This is useful in the analysis of digital mammogram images because the image is divided into various components and areas of interest. This section will present a survey of illustrative examples of image segmentation. There are numerous methods for image segmentation; among these are two basic types of image segmentation algorithms. This first group includes methods which attempt to differentiate regions of the image by locating the boundary or edges between different regions. These methods are referred to as edge based techniques. The second group of algorithms is those which attempt to

divide the image into segments by grouping individual pixels into various regions. These are called region based.

3.4.1 Edge Based Image Segmentation

Edge based segmentation methods attempt to define the regions of an image by locating the edges which form the boundaries for the regions. Figure 3.13 shows an image with simple regions. Using the Laplacian kernel in section 3.3.4, the edges in Figure 3.13 are detected and shown in Figure 3.14.

Figure 3.13: Example image with simple region

Figure 3.14: Edges outlining regions in Figure 3.13

This operation works very well for segmenting the simple image. The benefits of this are that the method requires no training data or prior knowledge of the image contents. However, this method alone is not suitable for images with more complex regions.

3.4.2 Region Based Image Segmentation

Region based segmentation methods typically evaluate each pixel in the image and compare it to the other pixels to determine distinct groups. This often requires a training set of data or a classification scheme to cluster the pixels into distinct groups. The following is an example of region based image segmentation. Linnett et al. established that pixel intensity distributions in a given area of an image follow a Poisson distribution. The Poisson distribution is a discrete distribution which describes a number of events occurring independently and randomly in time. Occurrences of pixel intensities in an image are examples of these events. To use this determination in segmenting an image, one must first select training areas from the image. The next step is to compute histograms for each of the training areas. Then the histograms from the training areas are compared to histograms from the regions surrounding each pixel in the image. This operation yields a probability that a given pixel belongs in a given training region. The training area which yields the highest probably for a given pixel is considered the matching region of that pixel. Thus the pixel is classified into a distinct region. Figure 3.15 shows the results of performing this segmentation method on the example image in Figure 3.13.

This method performs well on the example image and more complex images. However, it requires training data, which is not always available. Note also that this method segments regions based solely on the histogram distribution of pixel intensities.

This does not ensure successful segmentation because regions may have a different appearance and content while displaying the same intensity distribution.

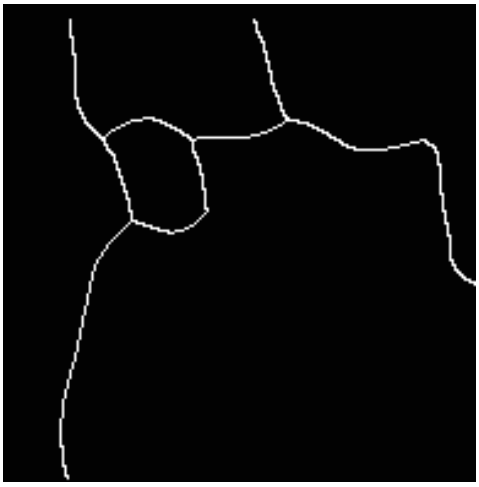


Figure 3.15: Poisson based region segmentation

3.5 Image Formats and Storage Methods

Digital images are typically compressed to conserve space. These compressions can be either lossless or lossy. Lossless compression preserves the information in the original image exactly, while lossy compression stores an approximation of the original image data. Medical imaging applications require lossless compressions. The images used in this research were lossless JPEG images, which are converted to PNG images for compatibility reasons to other software packages. Both are lossless.

Chapter 4

Digital Mammography and Computer Aided Screening

There has been an enormous amount of research in computer aided screening of digital mammogram images. It is represented in a broad array of published articles which describe the development, testing, and application of screening methods. Some research has developed specialized algorithms which solve specific challenges regarding computer aided screening. These include algorithms to enhance and/or segment images, detect/classify calcifications, or detect/classify masses. Other works involve combining these algorithms into more comprehensive systems, which may assist the radiologist or provide automatic detection of abnormalities. This chapter presents a brief survey of techniques developed for specified screenings and details both strengths and weaknesses.

4.1 Digital Mammogram Images

Digital mammograms are classified as either primary or secondary digital mammograms. Primary digital mammograms represent X-ray beams directly recorded as a digital image. Secondary digital mammograms are those that have been recorded onto film and then scanned into a digital format. The images used in this research are all secondary digital images.

> The digitization of mammographic images offers a number of potential advantages when compared to standard screen-film procedures. Digitized images may readily be stored within a computer system, allowing for improvements in data security and display flexibility and facilitating the rapid transmission of the digital image between clinicians at different stations. Additionally, image processing techniques may be applied to the digital mammogram in order to enhance the image or allow automated film reading. (Hutt)

These advantages are offset by a potential loss of resolution when changing from analog film to digital images.

A study by Karssemeijer, Frieling and Hendricks showed that the use of digitized mammogram images produced no significant reduction in the performance of radiologists in detecting abnormalities versus standard film methods. This study used digital mammogram images with a spatial resolution of 0.1mm per pixel. The film images used in this study have spatial resolutions between 0.042mm and 0.050mm per pixel. This is two to two and a half times the resolution of images used in the Karssemeijer study.

4.2 Preprocessing Methods

This section provides a brief survey of selected methods used primarily as preprocessing steps in the mammogram image screening processes. These methods do not directly locate or screen abnormalities. Instead, they enhance and improve the appearance or content of mammogram images to assist in the screening process.

4.2.1 Image Enhancement

In many systems, images are enhanced before the screening process. Enhancement procedures can amplify certain features or content in an image with the goal of improving screening effectiveness. For example, a screening system might enhance the image's edges before segmenting abnormalities like masses or calcifications. Image enhancement can also be used to diminish certain image features or content, such as noise reduction and/or background removal. Image enhancement may also include the separation of an image into various scales and/or the enhancement of a selected scale; which can amplify the features at a specified scale. Mammogram images may be enhanced before being

processed by other screening algorithms or the enhanced images may be prepared for direct viewing by radiologists.

In Veldkamp and Karssemeijer, an algorithm for normalizing local contrast in mammogram images is presented. They used a combination of methods for estimating high-frequency noise to create an adaptive algorithm for removing noise. It has been shown that this is a crucial step in locating microcalcifications. Microcalcifications occur in the high-frequency components of mammogram images and can be obscured by high-frequency noise. The normalization method is used as a preprocessing step to improve the performance of feature based microcalcification detection.

Kim et al. present a three step method for enhancing the appearance of mammogram images. First, film artifacts that may be misinterpreted as microcalcifications are removed. Second, gradient images are computed with first derivative operators. Finally, important features in the image are enhanced by adding, with adaptive weights, the gradient images from step two with the original image. This enhances image details and can improve the abnormality detection performance for microcalcification detection.

An image enhancement algorithm which improved mass detection was presented by Petrick et al. They developed a method for segmentation of mass regions in mammogram images using an adaptive density-weighted contrast enhancement filter with a Laplacian of a Gaussian for edge detection. The contrast enhancement filter enhances structures within the image so that the boundaries of objects in the image are more detectable. This is an effective preprocessing step; it segments masses from the image without a significant loss in the number of true masses.

4.2.2 Image Segmentation

Image segmentation algorithms have also been developed to assist in screening. Image segmentation algorithms serve two distinct primary purposes. First, the segmentation methods to separate image abnormalities from normal tissue are the most common. (Kupinski et al.; Huai Li et al). Second are the segmentation methods that divide the mammogram image into its large scale components: background, breast tissue, and pectoral region.

Many abnormality detection algorithms segment the breast tissue from the remainder of the image. One reason for segmenting the breast tissue from the image is this reduces the amount of image to be analyzed by follow-on steps in screening. These follow-on steps often include the computation of complex feature sets and a classification algorithm. Since these steps may be time-consuming, eliminating any unneeded analysis is beneficial.

Many detection algorithms locate abnormalities by comparing suspect features with their neighborhood. This is because most detectable mammogram abnormalities display properties that are different from their surroundings. However, these are only effective when localized areas can be segmented from the remainder of the image.

Segmentations algorithms are discussed in Section 3.4. Several mammogram image segmentation algorithms have been used in previous studies and they can be grouped into two categories: edge based and region based. Most algorithms for segmenting mammogram images are region based. They divide an image into distinct regions by classifying each pixel in the image as belonging to a unique region of the image. This is logical because mammogram images have unique, distinct regions. The

others first segment the mammogram into distinct regions by locating the edges/boundaries between the regions.

An algorithm that segments mammogram images into distinct components has been developed (Suckling et al.). This algorithm allows the image components to be analyzed separately. This segmentation algorithm uses nine statistical features from the image to train and segment the image. The features used are from first and second order histograms and include: mean, variance, skewness, kurtosis, entropy, contrast in the vertical direction at a distance of 3 pixels, contrast in the vertical direction at a distance of 7 pixels, correlation in the vertical direction at a distance of 3 pixels, and correlation in the vertical direction at a distance of 7 pixels. This is done with a multiple linked self organizing neural network to classify the feature statistics. This method provides visually good segmentation for the breast boundary, image background, and pectoral region. However, this technique requires training images to build the neural network classifier and may require retraining the system when using a different set of images. Neural networks are also computationally slow.

Some segmentation algorithms do simple thresholding to divide the image into different components. This works because the components of a mammogram image are often distinguished by their relative intensities and not their shape or texture. One difficulty with this type of system is to determine the proper threshold. Breast tissue is segmented from its background by locating a peak in the average contrast histogram, which is the ratio of the total contrast histogram to the standard gray level frequency histogram (H. D. Li et al.). This search for the peak occurs in the fixed intensity range of 70-120 with 8 bit grayscale images. This range was empirically chosen and is correct for

the image sets they used. The peak level is used to identify the boundary between the breast and its background.

4.3 Abnormality Detection Techniques

Another group of mammogram image analysis methods can be classified as abnormality detection techniques. These methods attempt to locate specific abnormalities, such as microcalcifications, mass lesions, or architectural distortions, within mammogram images. These methods draw from the field of image processing to locate abnormalities in mammograms.

4.3.1 Microcalcification Detection

Because microcalcifications are often an early sign of breast cancer, many researchers have developed methods to locate them in mammogram images. Microcalcification detection is not part of this research; however, these techniques are useful in other areas of mammogram image analysis.

Strickland and Hahn present a method for detecting microcalcifications which could also be classified as an image enhancement procedure. However, their method yields sufficiently refined microcalcification detections so that it is a standalone process. Their method detects microcalcification boundaries by decomposing the image into high and low frequency components using a wavelet transform. The high frequency components are dilated and then an inverse transform is performed. This enhances the microcalcifications to the extent that they may be segmented by simple thresholding. This detection performance is dependant upon the proper selection of a wavelet basis.

An example of a feature based method for locating microcalcifications is presented in (Dhawan et al). Their method uses a genetic algorithm to select an optimal

subset of feature statistics which is further used for the classification of microcalcification pixels. Then a feed-forward back-propagation neural network is used to classify the feature statistics.

Another method combines techniques similar to the previous two. It does a two step algorithm for locating microcalcifications in digital mammogram images (Yu and Guan). The first step segments potential microcalcification pixels using a combination of wavelet features and gray level feature statistics. Then individual pixels are grouped by spatial connectivity into distinct microcalcifications. The second step considers the grouped pixels as single microcalcifications and produces 31 feature statistics from each group. In each step a multilayer feed-forward neural network is used to classify the features.

4.3.2 Mass Detection

There are several algorithms for the detection of mass lesions that can be categorized. One category includes algorithms for locating potential masses. This type algorithm seeks to locate as high a percentage of tumors as possible and often returns a high number of false positives. This algorithm is often used in conjunction with another that classifies the possible mass location as benign or malignant. Methods that locate potential sites use a variety of image processing techniques.

A method is presented for locating masses in mammogram images by comparing the structures of the breast in the corresponding left and right mammogram images (Stamatakis et al.). They threshold the images based on local intensity and collect feature statistics from the remaining pixels. Then these statistics are compared to the mammogram image pairs to locate abnormalities. This method is useful because

radiologists often compare mammograms of the left and right breast for signs of abnormalities in either image. However, it requires a very accurate registration of the image components to permit the left and right images to be compared. For this reason, most researchers implement methods that consider each image separately when screening for masses.

Template matching based algorithms are also used to screen for masses in mammogram images. Morrison and Linnett present a model based approach to find objects in digital mammogram images. They use a bivariate Gaussian to model microcalcifications and a two dimensional sech function to model masses. Their method provides both a segmented region containing the mass and the center of the mass. Their method was tested on both simulated and a limited number of actual mammogram images.

Brake and Karssemeijer also used a template matching method for locating masses in mammogram images. They correlate areas from mammogram images with a sphere projected onto a two dimensional surface. This yields areas that are likely to contain masses. They compare their method with a statistical analysis of gradient orientation maps. In addition, they examined the convolution of a mammogram image with the Laplacian of a Gaussian. In their research, the template matching technique yielded the best results. They also investigated multiscale enhancements to the technique, but found those to produce little improvement.

Other methods for locating and classifying mammogram abnormalities are based on collecting feature sets from areas of the image. Feature sets consist of statistics based on regions of the image. The feature sets can be compared to other areas of the image or evaluated for content. The strength of these methods is to allow complex image data to be

condensed into a relatively small number of parameters. These parameters can be more efficiently interpreted than an entire image. Furthermore, a small set of feature statistics simplifies its use with a classifier, e.g. a neural network system, versus the entire image. For this method to work properly the features must be selected properly and correct weights must be determined (Brzakovic et al.; Kallergi et al.; Kobatake et al.).

Researchers often combine methods to locate masses in images. A template matching, edge enhancement or thresholding step will be applied to segment potential sites. Then, feature statistics taken from those sites are used with a classifier to refine the site selections. For example, Kobatake et al. present a two step system which uses an iris filter to locate generally rounded areas in the image. Then the process uses a set of features to distinguish between malignant and benign tumors. Their method correctly locates 90.4% of the malignant tumors in the images tested with 1.30 false positives per image. The 9.6% of tumors missed were largely due to sensitivity in the first processing step. Their results are significant; however, the first step should locate suspicious areas almost flawlessly.

Chapter 5

Segmenting the Mammogram Image

Accurate segmentation of mammograms can lead to improved computer assisted screening for breast cancer. The automatic segmentation of digital mammograms presented in this chapter accurately segments a mammogram into breast parenchyma, pectoral muscle, subcutaneous fat layer, and image background.

5.1 Importance of Proper Image Segmentation

“The female breast is composed of two major components: adipose tissue and fibroglandular tissue” (Sivaramakrishna et al.). Digital mammograms display both tissues. Figure 5.1 shows these components.

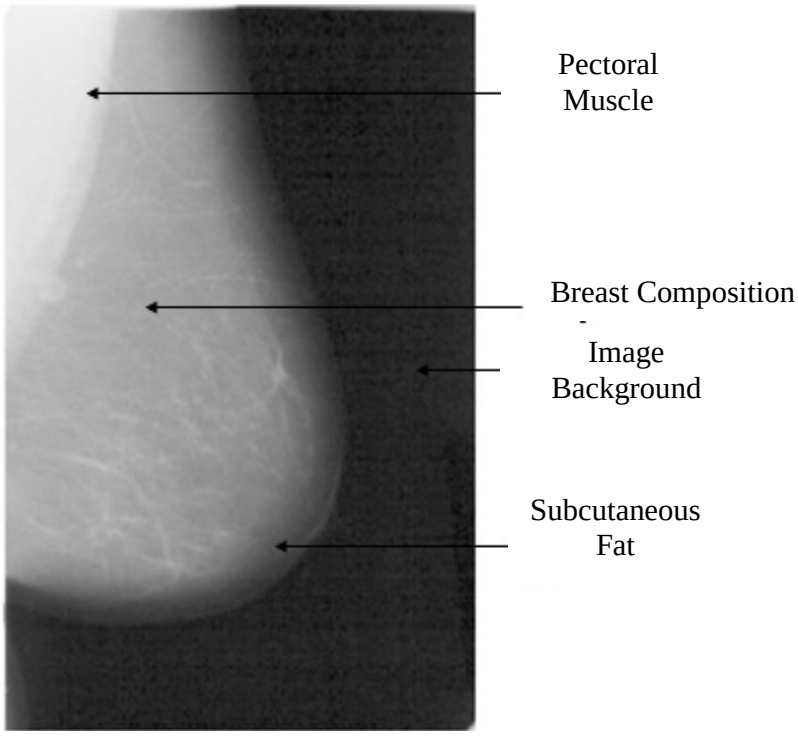


Figure 5.1: Mammogram image components

However, mammograms also display other components, which are often unnecessary or even detrimental to the diagnostic process in computer assisted screening. The mammogram image background does not need to be searched for signs of breast cancer. The examination of the black image background slows screening and can even produce false positives. We have developed a new method to automatically segment mammograms into breast parenchyma, pectoral muscle, subcutaneous fat layer, and black image background. Segmentation of breast tissue is an important pre-processing step that can be useful to other computer aided screening techniques. Many of these techniques first segment the breast tissue before attempting to locate possible signs of cancer, like masses or microcalcifications.

5.2 The Image Segmentation Algorithm

The most significant factor when visually segmenting a mammogram into major components is the relative intensity of each grayscale component. Texture, shape, and size are also factors, but the intensity of areas in the mammogram image is the best determinant of tissue type. Denser tissues appear brighter in mammogram images.

The background is the darkest portion of the image, the subcutaneous fat layer is slightly brighter, the parenchyma is even brighter and the pectoral muscle appears brightest in the image. In some cases simply dividing the image's grayscale intensities into four evenly spaced bands and then thresholding according to those bands can correctly segment an image. In an image with 256 grayscale levels, the four evenly spaced bands would be as follows: [0-63], [64 – 127], [128 – 191], and [192 – 255]. Figure 5.2 shows the mammogram from Figure 5.1 after this simple procedure.

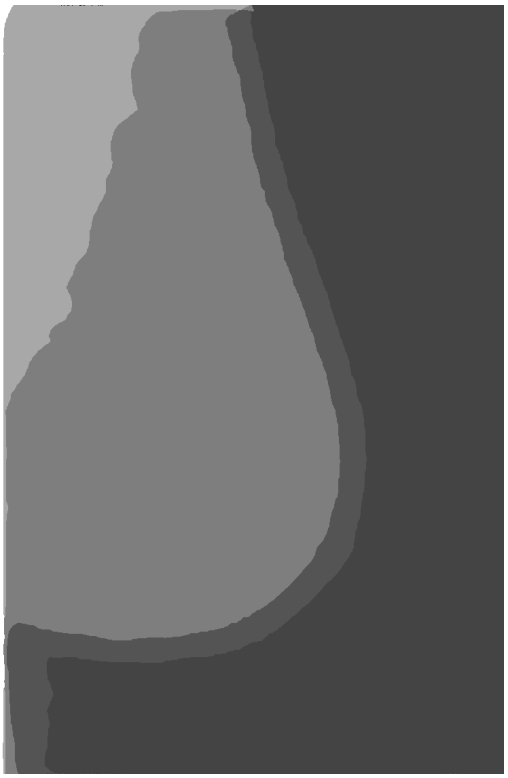


Figure 5.2: Segmented mammogram image

Figure 5.3 is the histogram of the image in Figure 5.1. It has four significant intensity bands, which correspond well with the four major image segments: breast parenchyma, pectoral muscle, subcutaneous fat layer and black image background.

Every mammogram image we examined displayed a unique but similar intensity distribution. However, the boundaries between the components of a mammogram image usually do not accurately correspond with the simple intensity bands used for the segmentation shown in Figure 5.2. A more sophisticated technique is required to determine the bands. In this chapter, we present an adaptive technique to accomplish the desired segmentation.

Our technique dynamically determines the proper intensity boundaries. Once the desired intensity ranges have been determined, pixels in the image can be grouped, resulting in an efficient and effective segmentation of the mammogram. Note the vertical

separations in the histogram in Figure 5.3; these represent the intensity boundaries between the various components of the breast. If these boundaries are properly located they can be used to segment a mammogram image.

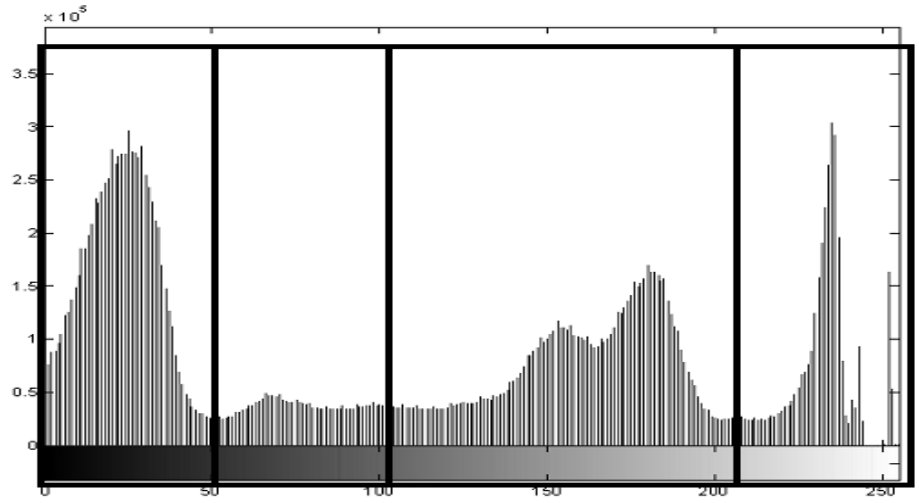


Figure 5.3: Mammogram image histogram with four bands marked

5.2.1 Histogram Based Segmentation

Other techniques for segmenting mammograms rely solely on its gray-level histogram. Some techniques use the spatial information, like shape and texture in a mammogram image, to segment the image. The technique presented here uses a combination of both intensity and spatial information. The first step in this algorithm uses the histogram to locate the intensity boundaries in a mammogram image.

Histograms are simple, fast to compute, and capture the intensity content of the image. (The area used to compute the histogram can be varied to change the scale of features being detected, i.e. segmented.) Our technique uses the histogram to form a feature vector and then a distance measure to segment the image into components. This is much faster than methods that utilize more complex methods, such as a neural network; this speed-up is significant. Also, this method requires no training data set.

The component boundaries of the mammogram image are characterized by changes in the pixel intensities in the local area. For example, the boundary between the pectoral muscle and the parenchyma will often be shown by a change in intensity of 30%-60%. The boundary between image background and breast tissue displays a similar change, but the boundary between subcutaneous fat and parenchyma is more subtle. Our method marks/highlights the areas of the image which display relatively high areas of intensity change. These areas are used to determine the specific intensity bands of the image and the desired segmentation.

5.2.2 The Histogram Difference Method

To measure intensity change, a histogram is formed for an area in the mammogram by summing the occurrence of each pixel's intensity in that area. The images used in this dissertation were originally 12-bit grayscale images, but have been converted to 8-bit grayscale images for use in our segmentation step.

The histogram for an image or sub-image can be viewed as a vector. In this case, the histograms have 256 values, so the resulting vector has 256 dimensions. A normalized dot product of two vectors will yield the cosine of the angle between the vectors in multi-dimensional space. This gives us a convenient and effective method for comparing two histograms. The dot product for two vectors $\vec{v} = (v_1, v_2, \cdots, v_n)$ and $\vec{u} = (u_1, u_2, \cdots, u_n)$ is defined as:

$$\vec{v} \bullet \vec{u} \equiv \sum_{i=1}^{n} (v_i * u_i) . \tag{5.1}$$

The angle θ between $\vec{v}$ and $\vec{u}$ can be determined as:

$$\cos(\theta) \equiv \frac{\vec{v} \bullet \vec{u}}{|\vec{v}| * |\vec{u}|}, \tag{5.2}$$

where $|\vec{v}| = \sqrt{\sum_{i=1}^{n} (v_i)^2}$ and $|\vec{u}| = \sqrt{\sum_{i=1}^{n} (u_i)^2}$.

The cosine of θ provides a measure of the "distance" between two vectors and is easy to calculate. This angle is viewed as a normalized dot product; it is used as a measure of the difference between histogram pairs. If two histograms are identical their normalized dot product is one, if orthogonal it is zero. With this, we developed a method to classify parts of a mammogram as the difference between combinations of histograms.

One could also use a variety of statistical tests to compare histograms, such as a chi-square test or statistical correlation. Some researchers have shown that statistical distributions, like the Poisson distribution, can be used to compare histograms (Linnett et al). However, for this research, we chose the normalized dot product because it is simple and fast to compute and it yields a value between zero and one in all cases.

The measure of difference between histograms of neighboring sub-images provides an excellent method for detecting gradual changes in the optical densities of a mammogram. The boundaries between pectoral muscle and breast tissue, and breast tissue and image background, can be obscured by noise. But a comparison of the differences between histograms produces an accurate location for the boundaries between pectoral muscle, breast parenchyma, subcutaneous fat, and image background.

For this to work effectively, a sampling method is constructed. The original mammogram image is divided by rows and columns into sub-images. These sub-images may be either overlapping or non-overlapping. The sub-image size may be varied to obtain the desired resolution of the algorithm. To accurately determine boundaries between breast parenchyma, pectoral muscle, subcutaneous fat layer and image background, adjacent sub-images are used that proceed from left-to-right and top-to-

bottom across the image. Figure 5.4 shows the mammogram image from Figure 5.1 divided into sub-images. This is only an example; in actual use the number of sub-images is much greater and the size much smaller. The sampling procedure is followed by taking the normalized dot product between histograms of pairs of sub-images. This dot product is used to construct a new image that identifies fluctuations in the optical density for the original mammogram. The new image is scaled down. For example, if the original image is 5000 x 5000 pixels using a sub-image size of 10, the resultant new image will be 499 x 499 pixels because the last column has no sub-image to its right for comparison. Figure 5.5 displays the results of this histogram difference applied to Figure 5.1.

Figure 5.5 shows small changes in the image scanning from left-to-right and top-to-bottom, but does not delineate the desired boundaries between the pectoral muscle, the subcutaneous fat layer, and the image background. In Figure 5.5, the lighter pixels represent sharp changes in texture, which correspond to either small features in the mammogram or significant boundaries between segments. The darker areas in Figure 5.5 represent areas with little change.

The image in Figure 5.6 better illustrates the effectiveness of the histogram difference method. It shows the results of the histogram difference method using only four-value histograms, rather than 256-value histograms as in Figure 5.5. Using four-value histograms has the effect of only highlighting the most significant changes in the intensities between images. Figure 5.7 shows the histogram difference result with four-value histograms overlaid onto the mammogram image from which it was taken. Note that the proper boundaries for image components are marked by the histogram difference measure.

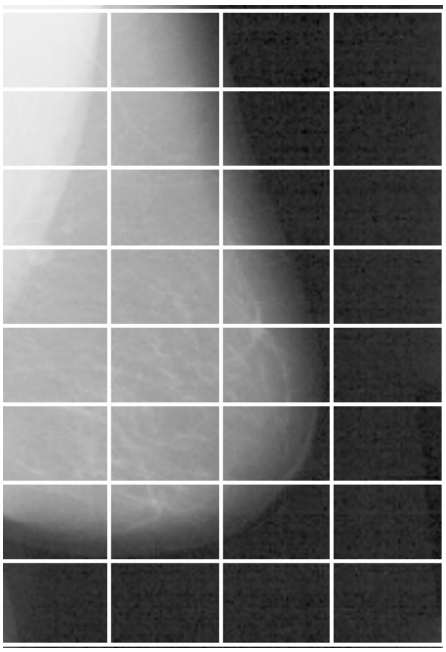


Figure 5.4: Mammogram image divided into sub-images

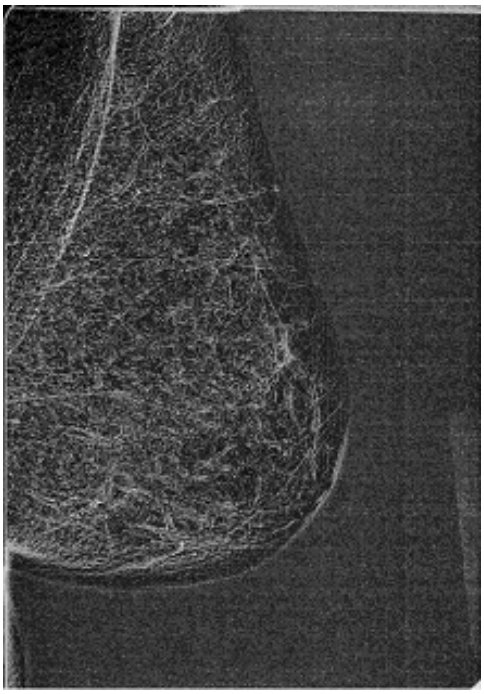


Figure 5.5: Histogram difference applied to mammogram image

Figure 5.6: Histogram difference result for four-value histograms

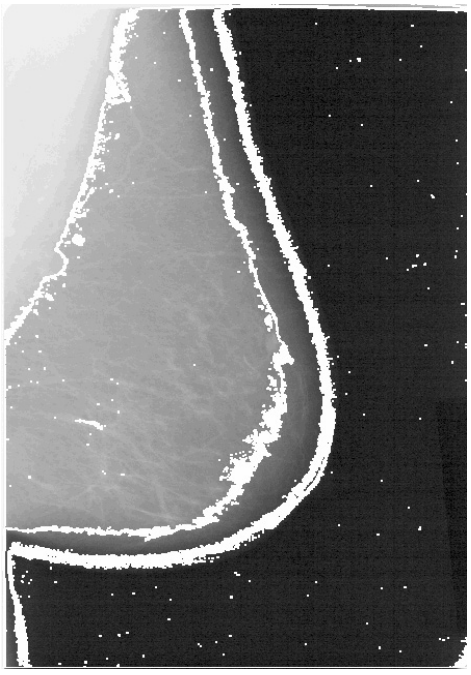


Figure 5.7: Overlay of histogram difference with mammogram image

The four-value histogram difference method works well in this case, but in other cases it may not accurately detect the proper boundaries. When the histograms of the mammogram sub-images are compressed from 256 values to 4 values, the boundaries

that will be detected are forced into following pixel intensity ranges [0-63], [64 – 127], [128 – 191], and [192 – 255]. These are the ranges used to produce the image segmentation in Figure 5.2. They work well for this image. However, not all mammogram images are the same. For example, Figure 5.8 is the mammogram image from Figure 5.1 with all the pixel intensities increased by 15%. This type of intensity variance is not uncommon between mammogram images, even between those in the same set.

Figure 5.8: Mammogram image with intensities increased by 15%

Figure 5.9 shows the result of applying the four-value histogram difference to the image in Figure 5.8. Note that there are several problems. Only three components were detected. The boundary between the pectoral muscle and the breast composition is not determined. Also, image noise in the background has been accentuated and will render

the results unusable in an automated environment. Finally, the boundary between the breast composition and the subcutaneous fat and the boundary between the subcutaneous fat and the image background are incorrect.

The same problem is encountered when the original image pixel intensities have been decreased by 15%. Figure 5.10 shows the result from applying the four-value histogram difference method to the decreased image. In this example, all the boundaries are incorrectly determined, and noise in the breast composition region renders the results unusable.

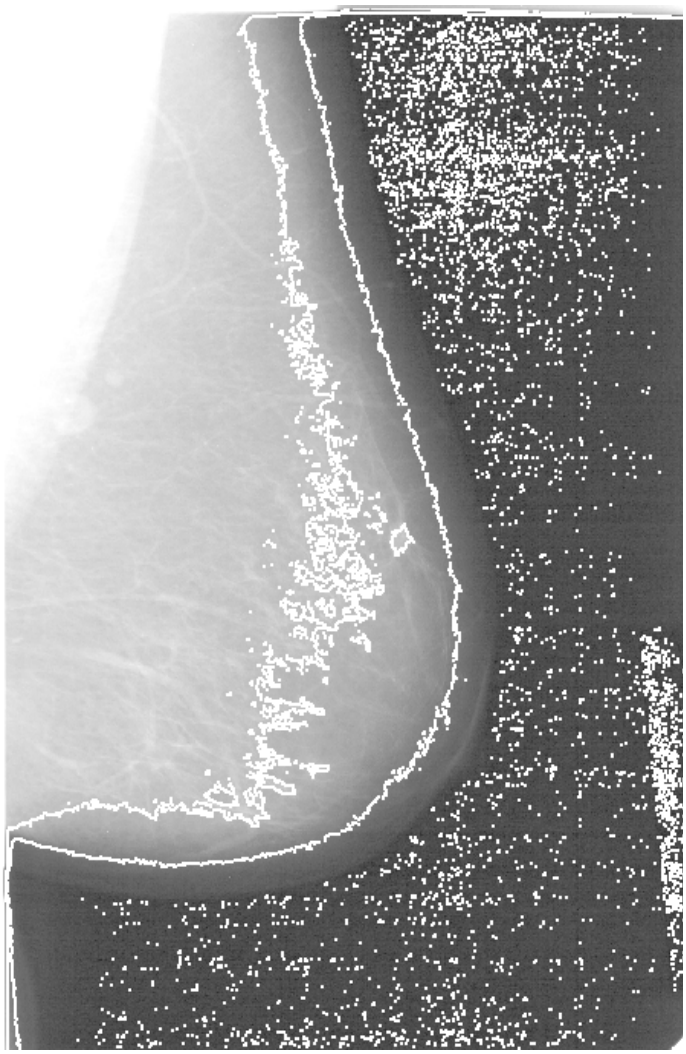


Figure 5.9: Overlay of histogram difference with mammogram increased by 15%

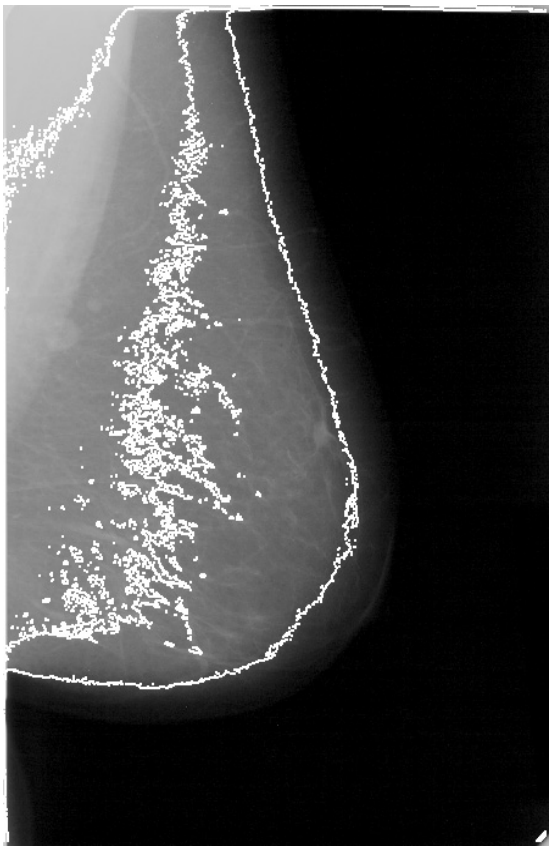


Figure 5.10: Overlay of histogram difference with mammogram decreased by 15%

The four-value histogram difference method is not sufficient to provide the boundaries between the image components. Therefore, we have developed a method to locate those boundaries using the full 256-value histogram difference method.

5.2.3 Finding Component Boundaries

The full 256-value histogram difference method result, as shown in Figure 5.5, produces a new image in which the pixel intensities corresponding to the amount of intensity change in that area. However, many of the areas of change represent small scale features in the image which are not significant boundaries between the image's components. To simplify segmentation, we apply an averaging filter, as discussed in Section 3.3.3. Figure 5.11 shows the mammogram image after a 30x30 averaging filter is

applied. Figure 5.12 shows the histogram difference method result on the image in Figure 5.11.

In Figure 5.12 the areas of change are more apparent than before and also more relevant to the significant boundaries between the image components. We locate the component boundaries by creating a weighted histogram of the original image with the results of the histogram difference measure acting as the weights. We construct a histogram of the original image, but instead of adding one for each pixel in the original image, our method adds the histogram difference value for that pixel in the image. All these values range between zero and one. This produces a modified histogram that emphasizes intensities which occur in areas of change in the image.

Then intensity values in the weighted histogram are normalized to eliminate any bias from larger boundary areas. Figure 5.13 is the histogram of the original image and Figure 5.14 is the weighted histogram created with the histogram difference image.

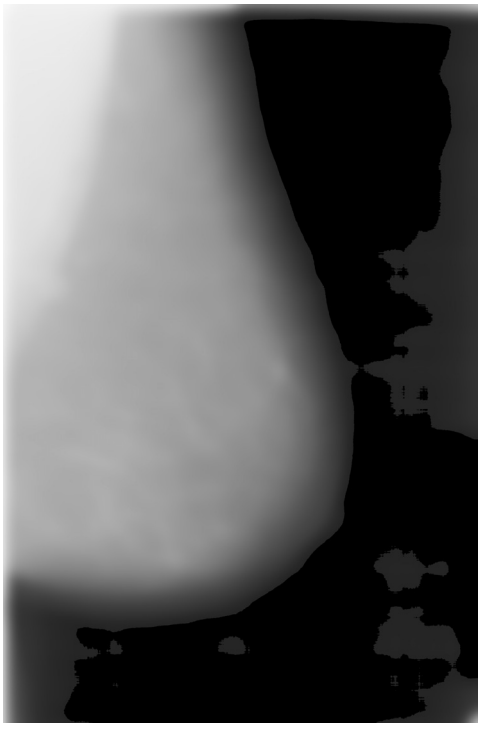


Figure 5.11: Mammogram image with small scale features removed

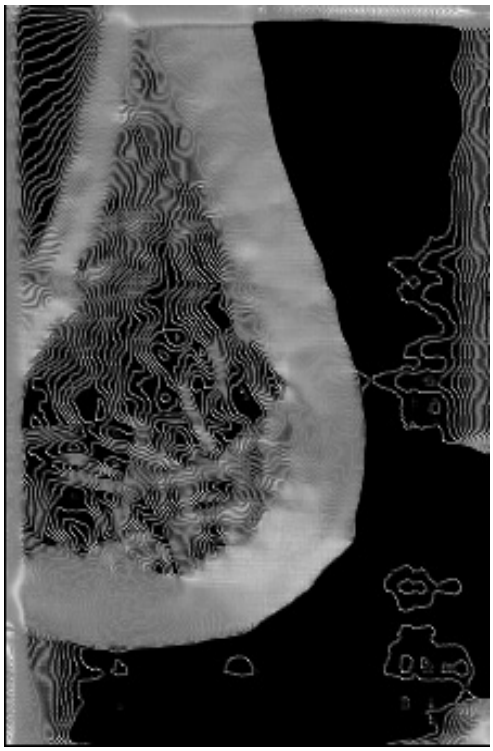


Figure 5.12: Histogram difference result on smoothed mammogram image

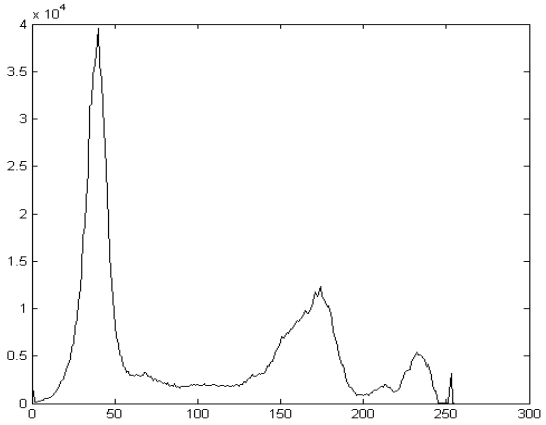


Figure 5.13: Histogram of original image

Using the weighted histogram in Figure 5.14, the desired boundaries are clearly identified. Figure 5.15 shows a simple first derivative approximation of the weighted histogram that has been thresholded at zero. The local maxima from this plot are the locations of the component boundaries. Figure 5.16 shows the relationship between the weighted histogram and the histogram difference image. In Figure 5.16, the circled areas

in the weighted histogram are the boundaries between the image's components. The arrows show the location of the weighted histogram boundaries on the histogram difference image.

By using the weighted histogram, we correctly locate the proper intensity boundaries for the image components and segment the original image. Also, the adjusted images shown in Figures 5.8-5.10, which were not correctly segmented by the four-value histogram difference method, are also correctly segmented. These images were the same as the original image, except that their pixel values had been changed by 15%. Therefore, the weighted histograms were virtually the same. The only difference was that they were shifted to the left or right by 15%.

The boundaries circled in Figure 5.15 represent pixel intensities that correspond to the boundaries of the image components. In this example, the intensity boundaries are 45, 81, 188 and 244. Thus, pixels with values from 0 to 45 were image background. Pixels with values 45-81 were subcutaneous fat. Pixels with values 81 to 188 were breast composition, and pixels with values 188 to 244 were the pectoral muscle.

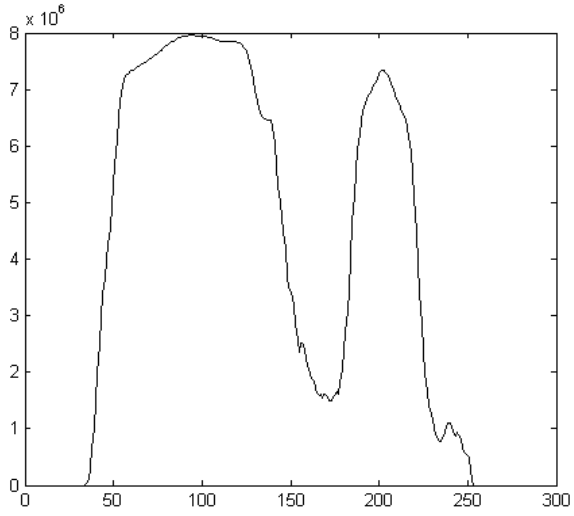


Figure 5.14: Weighted histogram of original image

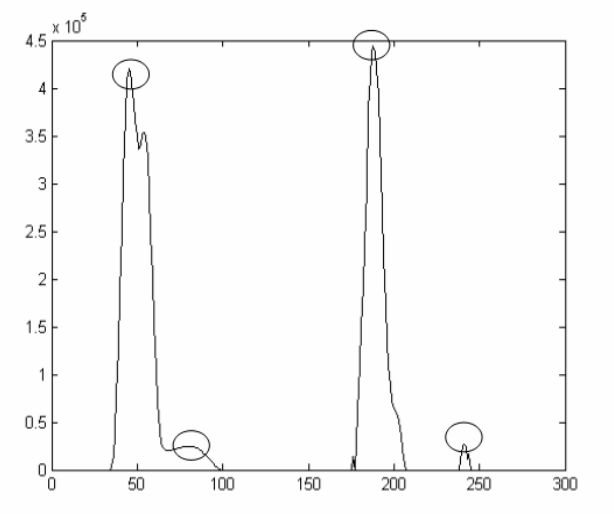


Figure 5.15: Location of component boundaries (circled)

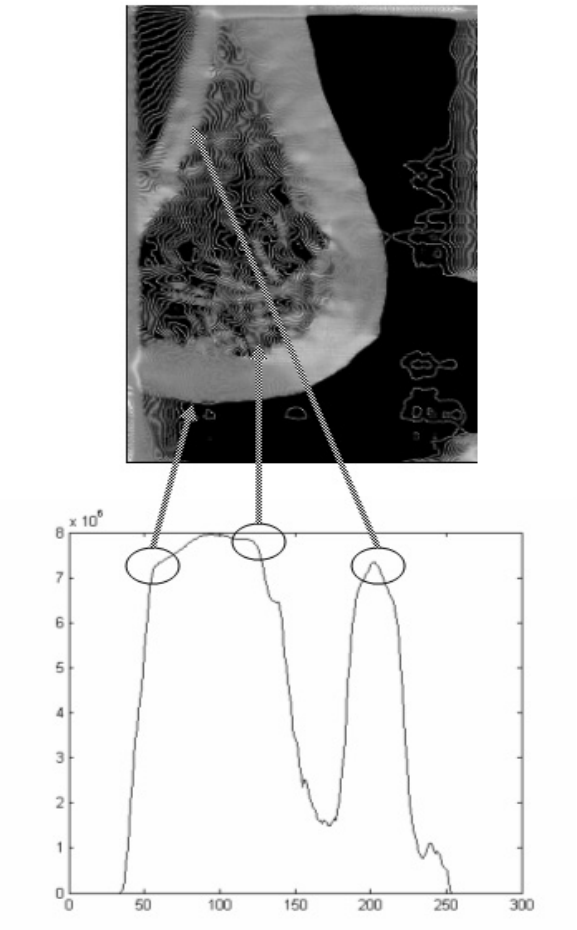


Figure 5.16: Relationship between weighted histogram and histogram difference image

5.2.4 Adding Spatial Context

The histogram difference method presented in the previous section produces the boundaries for the mammogram image components. However, those boundaries are in the form of pixel intensities. To segment the image into distinct components, the pixel intensities must be applied to the mammogram image.

Using the intensity boundaries, we threshold the original mammogram image. This procedure produces the image in Figure 5.18. Note that there are some misclassified pixels. These pixels are spatially located inside a component but their intensity values are outside of the detected range for that component. For proper segmentation and correct analysis, they should be changed to match the component to which they belong. This is easily accomplished by performing an erode operation on the image. The erode operation checks each pixel in the image and if a pixel is different from four of the eight which are its nearest neighbors, it is changed to match those four. In other words, if the pixel stands out from its background it is changed to match the background. This eliminates the misclassified pixels from the segmentation image.

Figure 5.18 shows the image from Figure 5.17 after one erode operation. Many of the misclassified pixels have been corrected, but some remain. To correct this, the erode procedure is repeated multiple times. However, this repetition of the erode procedure introduces bias into our image. This bias comes from the order in which the pixels are examined. If we start eroding misclassified pixels from the top-left corner of the image, then those pixels dominate the image and the results are unusable. The bias resulting from performing ten erode operations of Figure 5.17 can be seen in Figure 5.19. To eliminate this bias, we create a random ordering of the pixels in the image, and then we erode the image according to that random ordering. This produces a segmented image

free of positional bias. Figure 5.20 shows the properly segmented image, created by performing the randomized erode ten times on the image in Figure 5.17.

The image is now properly segmented. However, one step remains before the segmented components can be processed separately. The segmented areas need to be grouped into separate clusters and extracted into different images. This will allow them to be processed separately. An algorithm for creating groups of adjoining pixels of like color is presented in Chapter 7. The algorithm groups and extracts the segmented components in Figure 5.20. The extracted groups are shown in Figure 5.21 and represent pectoral muscle (top left), breast parenchyma (top right), subcutaneous fat (bottom left), and image background (bottom right).

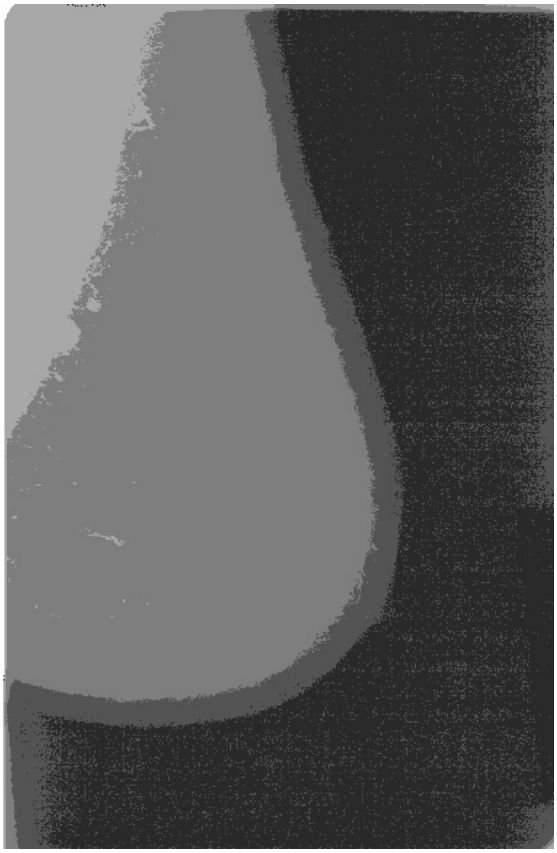


Figure 5.15: Mammogram image thresholded according to intensity boundaries

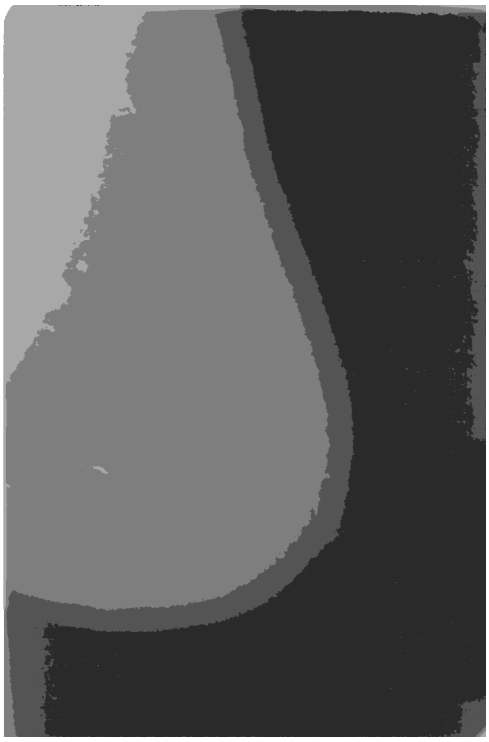


Figure 5.16: Segmented image with reduced misclassified pixels

Figure 5.17: Image segmentation with error

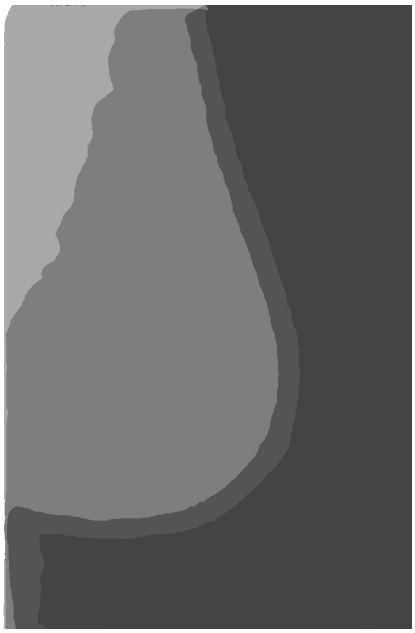


Figure 5.18: Correct image segmentation

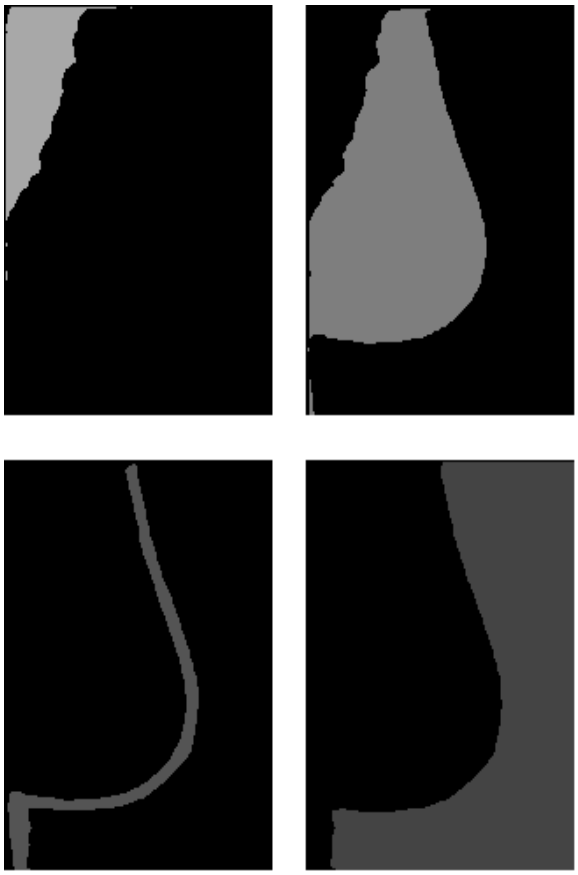


Figure 5.19: Image components extracted into separate images

5.2.5 Segmenting Dense Mammogram Images

Some mammograms display dense breast tissue which can make them difficult to segment and screen. Figure 5.22 is a mammogram image displaying dense fibroglandular tissue. Figure 5.23 shows the segmentation of the image in Figure 5.22. In this example the dense area of fibroglandular tissue is segmented separate from the breast parenchyma. This allows mammogram screening algorithms to analyze the segments separately.

Figure 5.20: Mammogram image with dense fibroglandular tissue

Figure 5.21: Segmented image with dense fibroglandular tissue

5.3 Segmentation Process Summary

The improved segmentation algorithm for mammograms presented in this chapter correctly segments mammograms into breast parenchyma, pectoral muscle, subcutaneous fat layer and image background. Our method is suitable for segmenting a broad set of images and, unlike other techniques, it requires no training data.

The steps of the segmentation process are:

1. The image is smoothed using an averaging filter to remove small scale features.
2. The histogram difference method is applied to create a new image.
3. A weighted histogram is created using the histogram difference image as weights.
4. The first derivative approximation of the weighted histogram provides the intensity boundaries of the image.
5. The image is thresholded according to the intensity boundaries.
6. The thresholded segmentation image is eroded multiple times to eliminate misclassified pixels.
7. Distinct components are extracted from the final segmentation image using the algorithm detailed in Chapter 7.

Chapter 6

Detecting Masses in Mammogram Images

Masses in the breast can be located in digital mammogram images by computationally analyzing various feature statistics from the image. Any algorithm used to analyze digital mammogram images can be both time-consuming and error-prone because many areas of these images appear to have features that are mass-like but not masses. Thus false positives are produced which detract from the effectiveness of the algorithm. In this chapter an efficient-straightforward algorithm to locate and record suspicious areas in a mammogram image is presented.

The algorithm presented has three steps. First, a template is used to identify and record suspicious pixels. A grouping of the suspicious pixels follows this, and finally the grouped suspicious pixels are retested with a template scaled to match the exact size of the potential mass.

6.1 Algorithms for Locating Masses in Digital Mammograms

Methods have already been published for the computer assisted detection of mass lesions in digital mammograms. These methods can be classified as either pixel or region based. Pixel based methods extract statistical features from each individual pixel in the mammogram image and use a classification scheme to identify and record pixels of interest. In some cases, a further examination could indicate if a mass represented by these pixels is benign or malignant. The other method is region based and it searches whole areas of the mammogram image for masses. This chapter presents a hybrid method for computer assisted screening of mammograms for masses. The hybrid method is a

combination of pixel and region based analysis. We employ a step, which narrows the focus of analysis from every pixel in the image to groups of pixels, which are areas in the image. These image areas are screened to determine if they have possible masses. Those with possible masses are then extracted and processed. This processing employs multiscale tests to refine the suspicious areas. This approach offers the potential for increased efficiency and reduced error for computer assisted screening of mammogram images.

Our technique uses a form of template matching at multiple scales to locate pixels in the image, which may be part of a mass. The resulting image is adaptively thresholded to a predetermined level of accuracy and then the remaining pixels are grouped together and extracted. This chapter includes a discussion of templates, template matching algorithms, and adaptive thresholding. Our technique was developed using a set of 156 images from 39 mammograms. A separate verification set of 160 mammogram images from 40 mammograms was used to validate the technique. Results from this testing are presented in Chapter 8.

6.2 Mass Detection Algorithm

6.2.1 A Template for Mammogram Masses

A mass in a mammogram image is either benign or malignant. To search for malignant masses we have developed a template that closely matches the properties of the masses in our developmental data. Masses tend to have a greater intensity than their neighboring regions. They are somewhat circular, although they display weak or fading boundaries with neighboring tissue. Figure 6.1 is an example of a cancerous mass identified inside the dotted circle.

We locate and identify objects in a mammogram image by comparing the image to the chosen template. The template is constructed from a part of another image that exhibits the visual and statistical properties of the objects being sought. Other authors have offered suitable templates (Morrison and Linnett; Brake and Karssemeijer). We thoroughly examined possible templates and chose the template that was most effective for the mammograms used in this research.

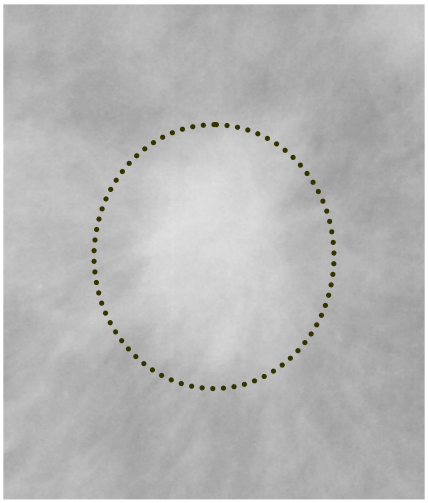


Figure 6.1: Example of a cancerous mass

To examine templates, we obtained a set of mammogram images with malignant masses. These come from the image set in (Heath, Bowyer and Kopans). These images are accompanied by pathology, which indicates the boundaries of the malignant masses. The images used were cases: 1118, 1134, 1156, 1159, 1160, 1163, 1166, 1174, 1203, 1212, 1217, 1222, 1224, 1229, 1236, 1252, 1262, 1403, 1417, 1467, 1486, 1520, 1557, 1587, 1589, 1592, 1620, 1622, 1642, 1671, 1693, 1700, 1701, 1720, 1726, 1790, 1896, 1899, and 1908. These mammogram images have a total of 80 malignant (cancerous) masses identified by radiologists and verified with pathology. Using these 80 masses, we recorded the performance of each template and made a selection based on these results.

We used a statistical correlation to compare a template with the actual mass. This comparison is detailed in the next section. Since, the size and location of each actual mass is known, a template of the same size can be generated and centered about the same location for comparison. Also, the effects of slight errors in comparisons, when the template is not centered exactly with the center of the mass or when the template is sized differently from the size of the mass can be viewed. Using this technique for comparison, four different templates were selected: a sphere projected onto 2-D, a 2-D hyperbolic secant function, a simple circle, and an actual malignant mass; shown in Figures 6.2-6.5. The sphere projected onto 2-D, which is shown in Figure 6.2, is generated with equation (6.1) as suggested by Brake and Karssemeijer. *D* is the diameter of the template.

$$F(x, y) = D^2 - x^2 - y^2 \qquad (6.1)$$

The 2-D hyperbolic secant (sech) function shown in Figure 6.3 is generated with equation 6.2 as suggested by Morrison and Linnett.

$$F(x, y) = \sec h(x + y) \qquad (6.2)$$

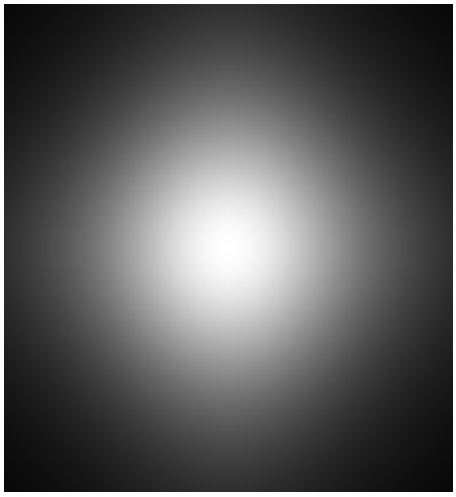


Figure 6.2: Projected sphere template

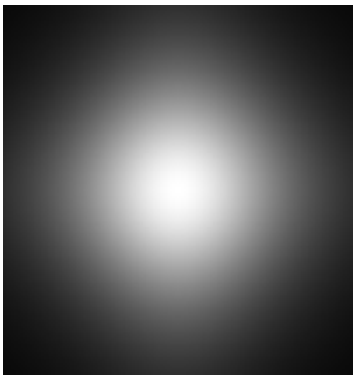


Figure 6.3: 2-D sech function template

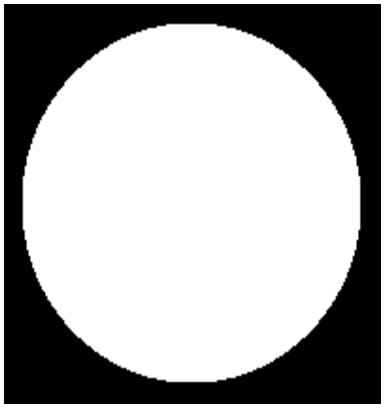


Figure 6.4: Simple circle template

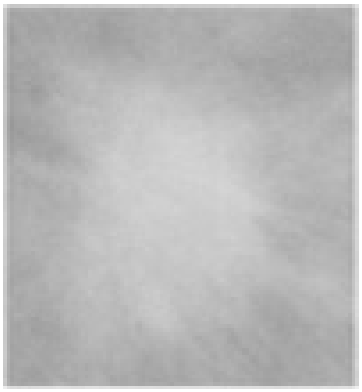


Figure 6.5: Malignant mass template

Masses in the breast are generally circular and symmetric. The templates in this example are also symmetric and circular; therefore, orientation errors are negligible. Every pixel in a mammogram image is included in the template matching process, i.e. the center of any possible mass will also be included. This eliminates any error from template

comparisons that have not been centered on a possible mass. These simplifications allow any effects the scale has in judging different templates to be included in its consideration for future use.

The primary measure for selection for the various templates is the template sensitivity to scale errors. When a template is compared to portions of the mammogram image, a specific sizing of the template will have to occur. This sizing of a template will probably be different from the size of the actual mass in the image; this sizing introduces scale errors. In other research, templates of multiple scales were used as part of the template matching process (Brake and Karssemeijer); it produced little benefit over a single scale. Therefore, when a template is selected and subsequently used to search the image, a single template, which will match masses of varying size, is desirable.

To evaluate the four chosen templates for suitability and scale sensitivity, each template is compared to all 80 of the actual masses. A template is generated that is the exact size of each mass and for each scale of the multiple scales being used. This permits consideration of the effectiveness of each template for masses of varying size.

Statistical correlation between the template and an actual mass yields a value between -1 and 1, where 1 indicates an exact match. Table 6.1 shows the average correlation between each template and the 80 masses and when the templates and the masses were the same size. The sech template returned the highest (best) correlations. This is best because the higher the correlations, the better the thresholded results, i.e. they can determine the correct-desired results better.

Figure 6.6 shows how the four templates correlate to the known masses when their sizes are scaled. This plot shows the results of correlating templates scaled from 0.1 to 2 times the size of the actual mass. The best performing template in this test is also the

sech template shown on Figure 6.6. It is consistently the highest (best) and has less sensitivity to scaling errors. Note: all templates perform poorly whenever the scale significantly differs from their size. In particular when the template size is several times smaller than the actual mass size, the results are unsatisfactory. However, when the template sizes are larger than the actual mass sizes the results degrade slowly and are useable. Therefore, to choose template size that is too large is better than to choose one that is too small. From this, the sech template is better than the other templates studied. The sech template is chosen for the remainder of this dissertation research.

Table 6.1: Average template to mass correlation values

Template	Average Correlation
Sech	0.7992
Tumor	0.6470
Circle	0.5480
Sphere	0.7502

6.2.2 Template Matching Methods

There are several ways to compare a template to a mammogram image. Some use convolution or a cross-correlation based method, which is very efficient with the use of the Fast Fourier Transform (FFT). However, we found that convolution methods with the sech template can become disordered by changes in local mammogram intensity. For example, a mass may appear significantly brighter than its neighbors' average intensity, but when compared to all other areas and tissues in the image, it may be darker than the average global intensity of the image. This mass would be missed using a convolution-based template matching algorithm. Figure 6.7 presents an example in which a mass in a local sub-image is correctly identified with a convolution method using the sech template. The left panel in Figure 6.7 is the actual image of the mass and the right panel is the result from the convolution thresholded to the appropriate level.

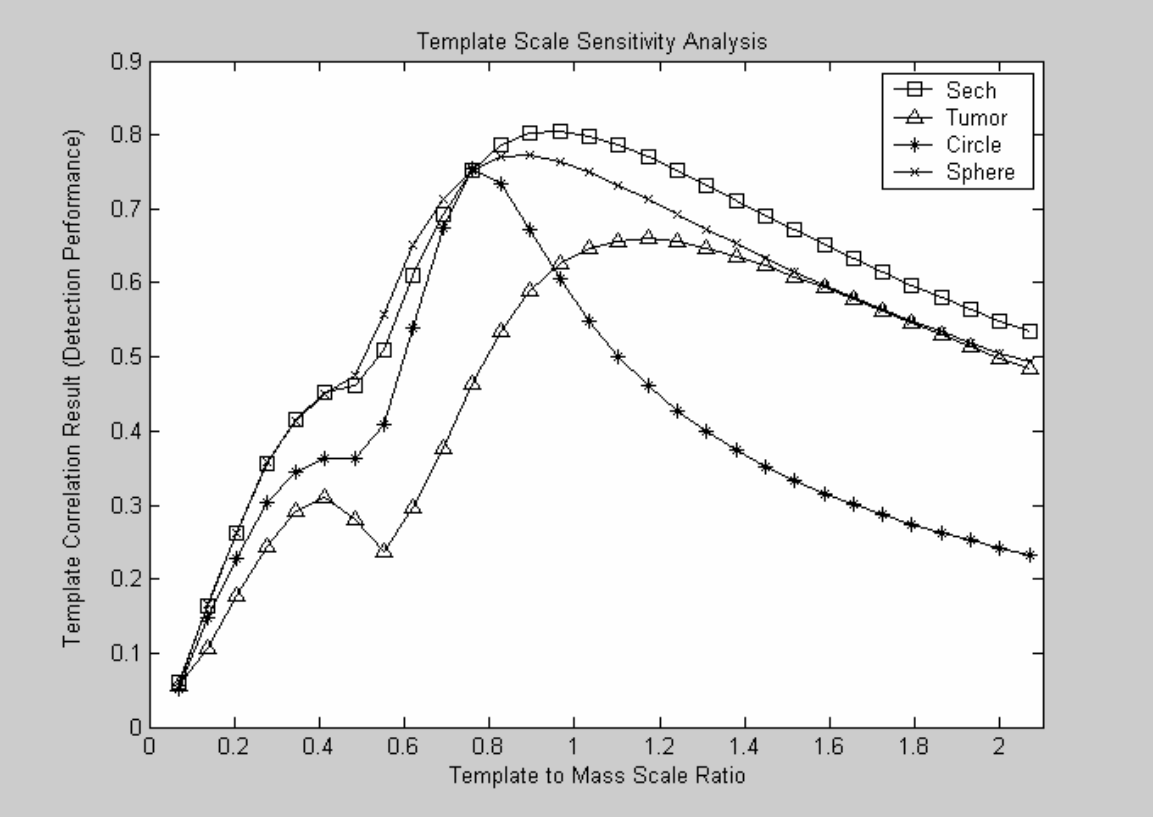


Figure 6.6: Template scale sensitivity

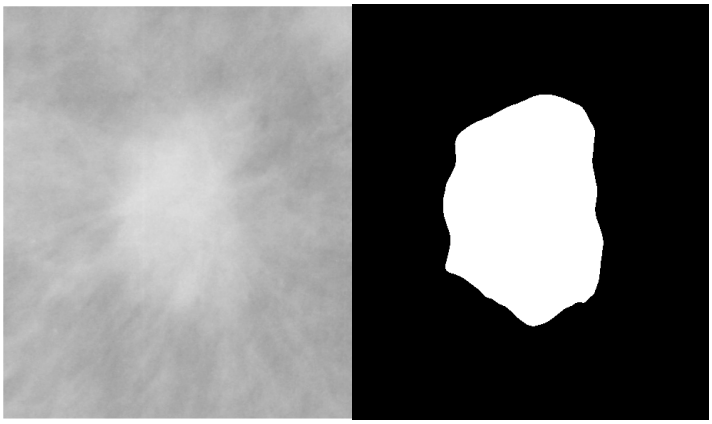


Figure 6.7: Mass (left) and convolution mass detection (right)

The results are not as effective as this example when the procedure is performed over the whole mammogram image. Figure 6.8 shows a mammogram image with a mass. Figure 6.9 shows the thresholded results of convolution over the whole image. The convolved results favor areas of high intensity, including the pectoral region of the

image. In this case the mass is not accurately located. Therefore, the convolution method is unsuitable as an effective mass location algorithm with this image.

An alternative would be the correlation between the template and each pixel in the image, where the sub-image surrounding each pixel is the same size as the template. The correlation-based template-matching algorithm is not biased by varying intensities and is sensitive to characteristics of the shape (Brake and Karssemeijer). The formula used for the correlation between the template and the original image is as follows:

$$cor(x, y) \equiv \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y}, \qquad (6.3)$$

$$\mathrm{cov}(x, y) \equiv \mu_{xy} - \mu_x \mu_y$$

$\mathrm{cov}(x, y)$ is the covariance of the template, x, and the sub-image, y, surrounding the pixel. σ_x and σ_y are the standard deviations of the template and the sub-image. μ_x and μ_y are the means of the template and sub-image. μ_{xy} is the mean of each pixel in the template multiplied by each pixel.

The result of applying the correlation to the image returns a new image, containing values between -1 and 1. The closer the correlation value is to 1, the greater the similarity between the template and the pixel being tested. That is, pixels producing a correlation value close to 1 are similar to the mass template. Figure 6.10 shows the results of applying this correlation to the image shown on Figure 6.8.

To find areas in the image, which are most likely to contain mass lesions, a threshold is used with the correlation values. In this example, pixels with correlation values of 0.75 or more have a high probability of being a mass lesion. Figure 6.11 presents the results of this thresholding with the image shown on Figure 6.10. The

threshold step reduces the number of pixel values to be considered and locates the part of the original image that has a high probability for mass lesions. In Figure 6.11, the white area indicates the mass location in the mammogram image.

Figure 6.8: Example mammogram image containing mass

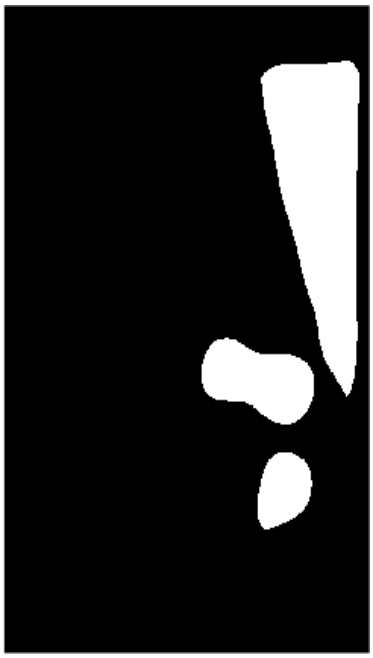


Figure 6.9: Results from convolution of template with the image in Figure 6.8

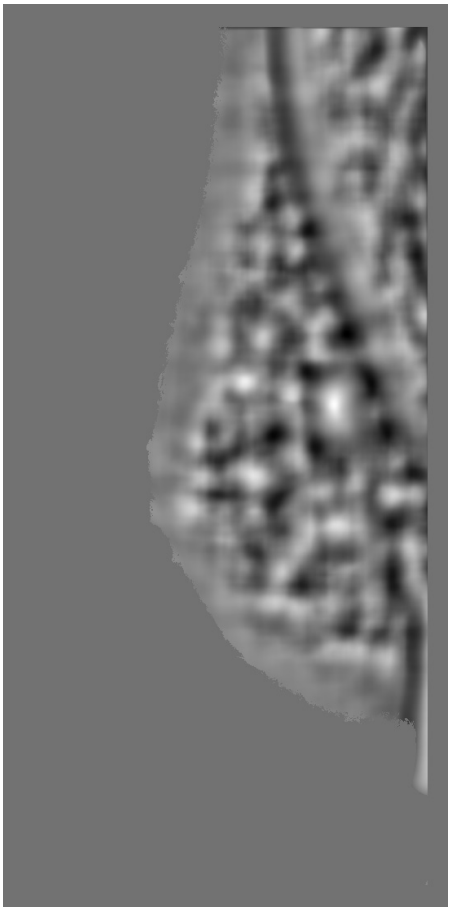


Figure 6.10: Results of correlation based template matching

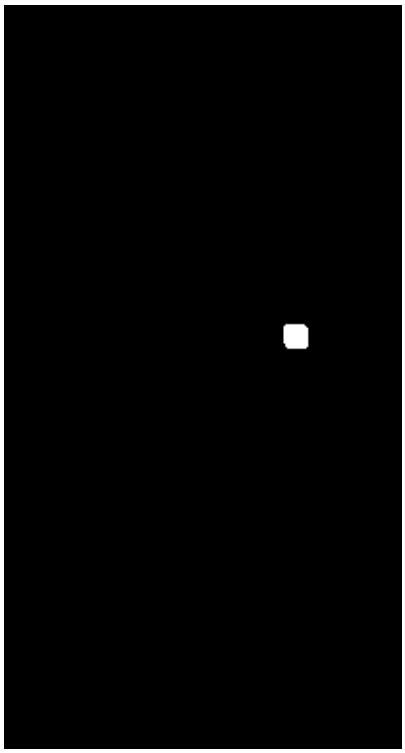


Figure 6.11: Thresholded and dilated results of correlation template matching

6.2.3 Clustering and Segmentation of Suspicious Areas

The example in the previous section demonstrates how to locate a mass in a mammogram image by performing a correlation followed by thresholding. The threshold in the previous example was selected manually. In actual use, a threshold would have a value lower than optimal. It is done so that few masses are missed by the analysis.

When a threshold value is reduced, more masses are usually located, but many of these are false positives. Figure 6.12 shows the result of thresholding the image on Figure 6.11 with a threshold value of 0.60. This produces several additional clusters of probable mass lesion pixels in the resultant image. There is only one actual mass.

To this point, the processing has been totally pixel based. Each pixel is essentially processed independently of the other pixels in the image. However, the pixels shown by the image in Figure 6.12 are parts of distinct clusters. Chapter 7 presents an algorithm for

clustering and extracting related pixel groups. This allows regions of similar pixels to be processed together. Figure 6.13 shows the four distinct pixel clusters in Figure 6.12. Each cluster is a potential mass and needs to be processed further and compared with its pathology.

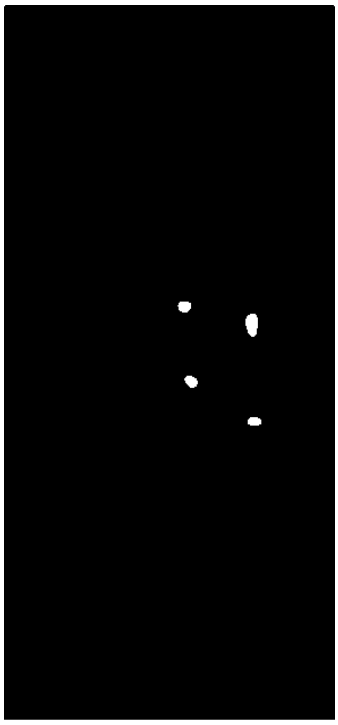


Figure 6.12: Correlation results thresholded with 0.60.

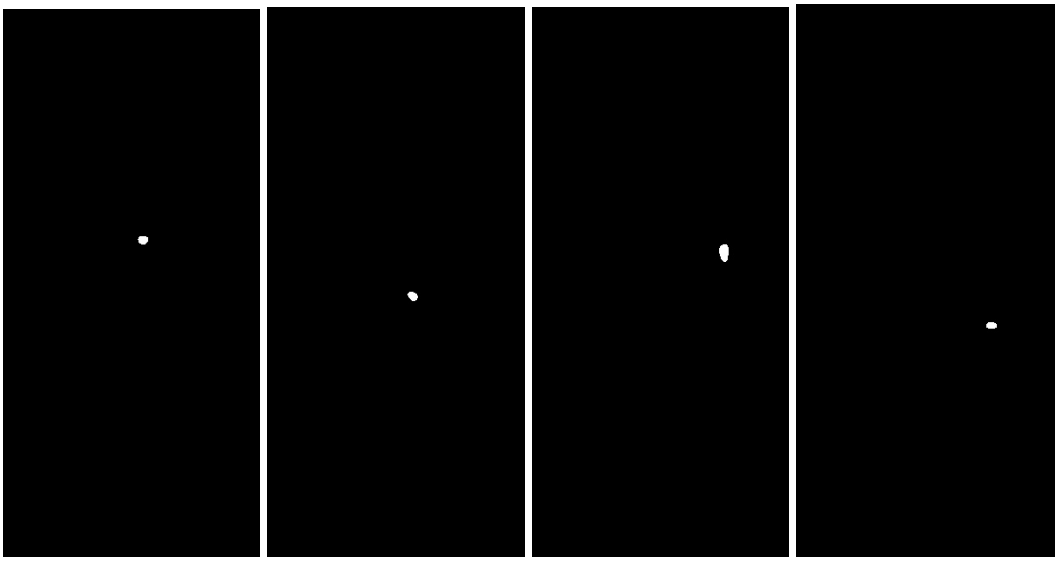


Figure 6.13: Extracted potential masses

6.2.4 Multi-Scale Enhancements

With the distinct pixel clusters, shown in Figure 6.13, a multiscale step is added into our mass detection algorithm. Section 6.2.1 discussed problems choosing a template that is a different size from the mass being studied. When the template is significantly larger or smaller than the mass of interest, it may not be detected. One solution for this problem is to test the entire image with different sizes of templates. However, this is time-consuming and ineffective.

Our solution is to do the template correlation on only the pixel clusters and use multiple-sized templates. This approach has the following advantages:

- The pixel clusters are only a small portion of the image and so, this multiscale testing is computationally feasible.
- The exact size of the pixel cluster has already been determined; therefore, a template can be generated for this size and used with the original image to eliminate any errors of scale.
- The exact location of the pixel cluster is known; therefore, this template can be centered in the cluster, rather than at non-central locations. This greatly speeds up the correlation.

To illustrate, the four possible masses shown Figure 6.13 were extracted from the mammogram shown on Figure 6.14. For each of these four possible masses a sech template is generated. Then the statistical correlation is computed between the correctly sized and centered template with the possible mass. Table 6.2 displays the results of this. The possible masses produced correlation values between 0.5448 and 0.8583. The actual mass has the highest (best) correlation, 0.8583. The other three possible masses were false positives. The results of this multi-scale process distinguished between the actual

mass (from pathology) and false positives. For this mammogram it is possible to determine and eliminate the false positives by using a higher threshold value. False positive errors can probably be reduced by this threshold addition.

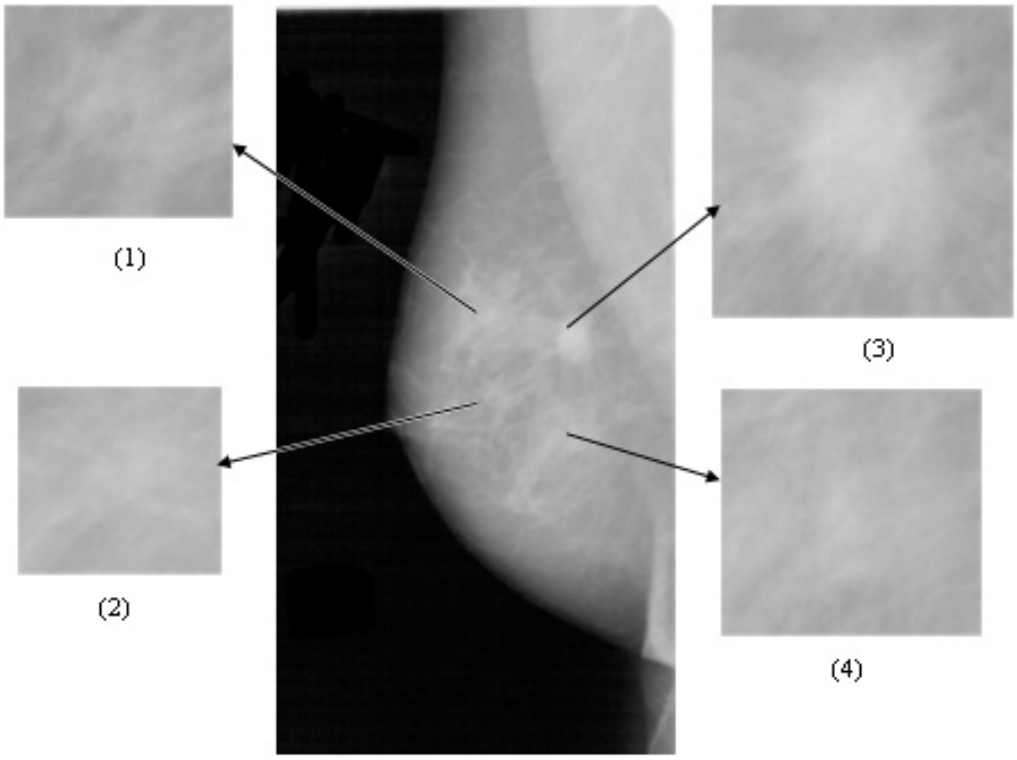


Figure 6.14: Possible masses from mammogram image

Table 6.2: Multiscale correlation results for possible masses

Object	Size (pixels^2)	Correlation to Template
(1)	30,000	0.6585
(2)	23,000	0.5448
(3) Actual Mass	**40,000**	**0.8583**
(4)	60,000	0.5919

6.3 Mass Location Summary

This new algorithm for locating masses in digital mammogram images follows:

1. The entire breast image is processed with a correlation-based template matching algorithm.
2. The result of this correlation is thresholded to determine the suspicious pixels.
3. These suspicious pixels are grouped into clusters and extracted.
4. The extracted clusters are each compared with a template that is constructed to be the same size as each cluster.
5. Then, each cluster is correlated with its template and the result of each correlation is thresholded again to eliminate false positives.

This algorithm can be used for mass detection or it can be combined with other steps to improve performance. For example, the segmentation algorithm in Chapter 5 can be used to eliminate background and other areas from consideration. This saves time and reduces errors when using with this mass detection algorithm. Also, the results of this mass detection algorithm are recorded as suspicious regions of the image. These regions can be subjected to further classification, which can also reduce error.

Chapter 8 presents the results from testing the mass detection algorithm. It also discusses how to choose the parameters to optimize the performance. These parameters include: template size, initial threshold, multiple scales used, and final threshold.

Chapter 7

Screening Algorithm Testing System

A mammogram screening algorithm-testing system was developed, a byproduct of this research. This chapter presents this system, which records and compares results from computer assisted screening mammography to their pathology. This technique provides detected suspicious areas from the template matching for comparison.

Many mass location algorithms produce results, which do not produce a simple "yes/no" answer for the location and malignancy of a tumor. Instead a gradient of "likelihood" is produced. One section of the image is determined to be more likely to contain a mass than other sections of the same image. One disadvantage of this is the system is more complex because a method for thresholding the results must be added and a proper threshold selected. After this, the threshold produces a "yes/no" answer for the more likely location of a possible mass. One advantage of this is that the threshold can be adjusted to produce almost any desired level of accuracy.

In order to set a desired level of accuracy for a mass screening system, various algorithm parameters must be determined by experimentally selecting "optimal" values. This experimentation is tedious and time-consuming particularly when the screening algorithm results are manually tabulated and compared to pathology. The method presented in this chapter provides a solution to this problem. It allows mass screening development time to be reduced and most of the results to be verified automatically. This method was used extensively in the development of the mass location algorithm presented in Chapter 6.

7.1 Image Filter Results

The mass detection algorithm presented in Chapter 6 produces an entirely new image with generated values in the place of the original pixel values. These new pixel values indicate a property this method is designed to promote. In our mass detection algorithm, the generated pixels are between 0 and 1. These pixels are correlation results between the selected template for a mass lesion and the other areas in the mammogram.

These correlation values are interpreted as the "likelihood" that each pixel is located at the center of a mass lesion. Figure 7.1 depicts a mammogram image and Figure 7.2 shows the result of the mass detection algorithm with this mammogram image. The area circled in Figure 7.2 is the known location of a malignant mass. The pixels in this area vary from 0.0 to 0.9; ideally, there would be a single pixel with a value 1 located at the center of the malignant mass. However, the algorithm is not perfect because there is no pixel with a value 1. Instead, a group of pixels in the area have values near 1. All these pixels with values near 1 are possible centers of the malignant mass. The other non-zero pixels present in the image are less likely to be the center of a mass and are probably image noise, normal breast features, or non-central areas of a mass.

Since the location of detected masses in Figure 7.2 is not precise, further analysis is required to produce a final result. This additional analysis begins with the resultant image being thresholded to eliminate results that are considered less significant. The resultant image is binary, i.e. all of the pixels are either zero or one. Zero indicates no mass, and one indicates a mass location. Figure 7.3 shows the result of this thresholding on Figure 7.2. For Figure 7.3, pixels with values over 0.75 on Figure 7.2 were assigned a value of one and all others zero. If the threshold is lowered, more detections will be produced, and if increased, fewer detections will be produced. This testing allows

thresholds as well as other system parameters to be varied and the results to be automatically applied for comparison. This procedure produces and can select an optimal threshold.

7.2 Clustering Pixel Groups

With the thresholded output from the image filter shown on Figure 7.3, the locations of detected masses are not completely determined. Each white pixel on Figure 7.3 represents a possible location for a mass; however, the pixels are clustered into distinct groups which represent single detections. These pixels in the distinct groups must be clustered and extracted as a single mass to be compared with the pathology.

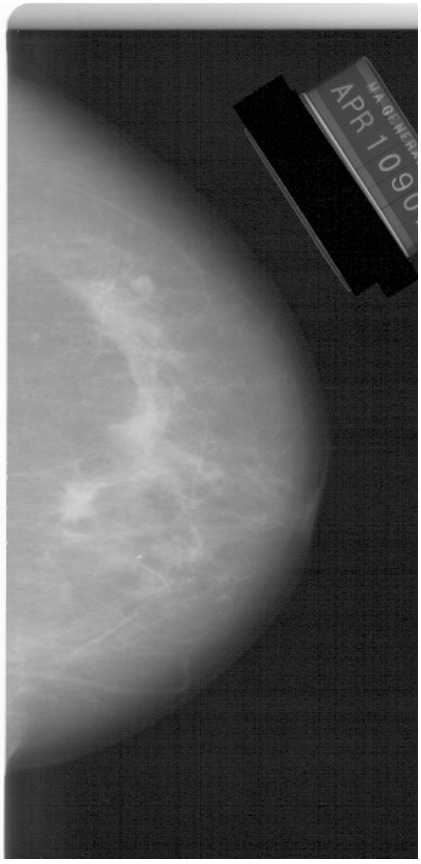


Figure 7.1: Mammogram image

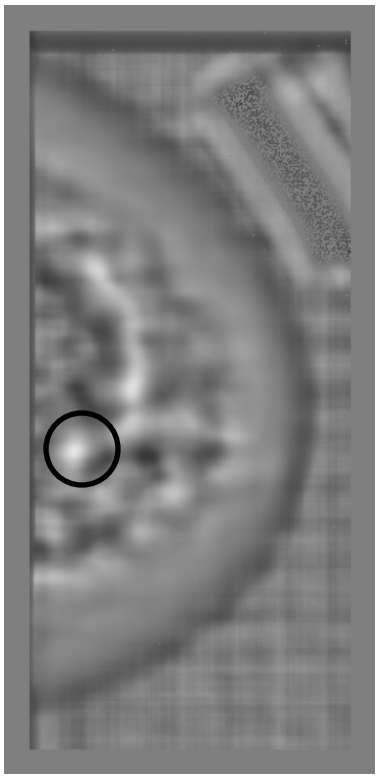


Figure 7.2: Results of mass detection on Figure 7.1

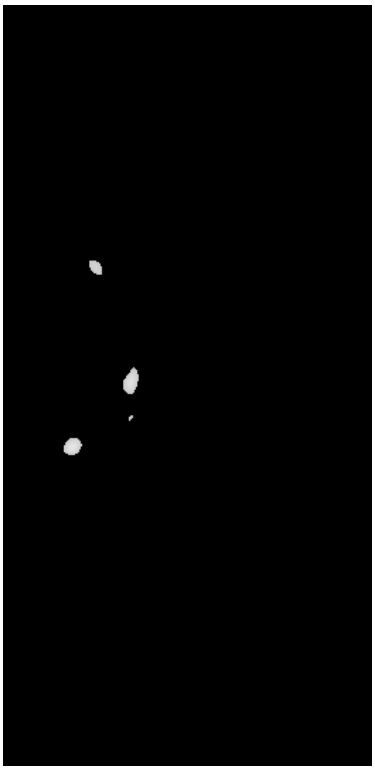


Figure 7.3: Thresholded image from Figure 7.2

Clustering the distinct pixel groups is not a trivial task. It requires checking every pixel in the image and clustering all the pixels of the same intensity that are nearest neighbors. Pixel *A* is a nearest neighbor of pixel *B* when pixel *A* is one of the eight pixels that immediately surround pixel *B*; Figure 7.4 presents this pixel and its eight nearest neighbor pixels. Figure 7.3 has four distinct groups of pixels and Figure 7.5 shows the four extracted groups with a separate image for each.

i-1, j-1	**i, j-1**	**i+1, j**
i-1, j	**i, j**	**i+1, j**
i-1, j+1	**i, j+1**	**i+1, j+1**

Figure 7.4: Pixel (i, j) and surrounding pixels

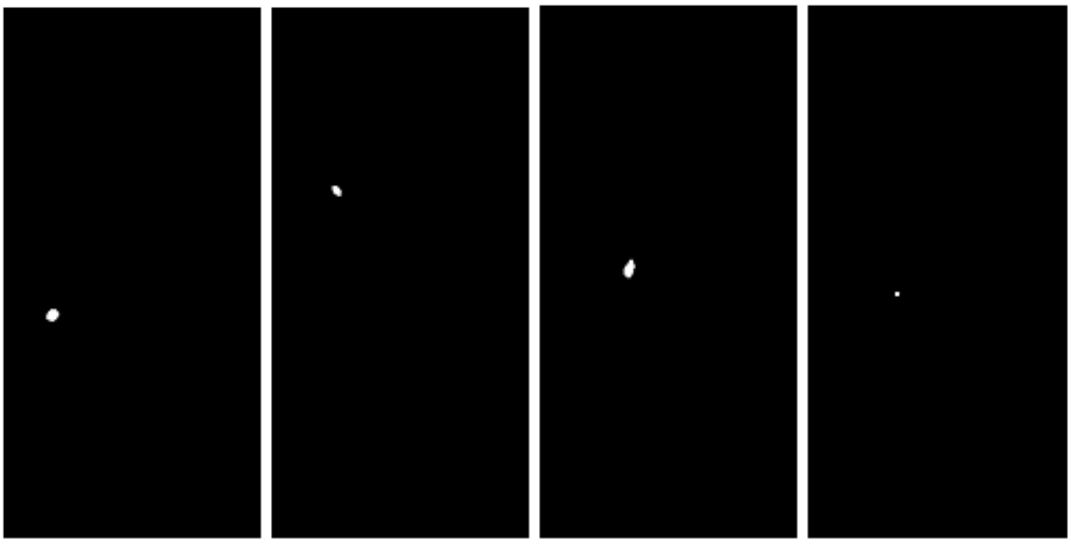


Figure 7.5: Four pixel clusters from Figure 7. 3

The algorithm that clusters and extracts the pixel groups in Figure 7.5 is part of the results of this dissertation research. The steps for this algorithm follow:

Pixel Clustering Steps

1. Two lists are created. The first is a pixel-list for all the pixels in the image. The second is an empty list, which will contain the clustered groups of pixels.
2. The first pixel in the pixel-list is removed from the list and processed. This is repeated until the pixel-list is empty. To process a pixel, each group in the group-list is checked. If the pixel being processed is a nearest neighbor of any of the pixels in one of the groups, it is added to that group. Each pixel will be added to only one group in the group-list.
3. Once the pixel-list is empty, all the pixels have been added to a group. However, in most mammogram processing, some of the groups in the group-list can still be clustered. Therefore, the next step requires that each group in the group-list be compared to the other groups. Those that contain nearest neighbors are combined (further clustered).
4. Finally, each of the groups in the group-list is formatted and stored as a separate image.

Figure 7.6 presents a diagram of this process.

Each mammogram images usually has more than a million pixels, and inefficient processing can easily produce impractical computational requirements. If *N* represents the number of pixels in an image; our algorithm for clustering groups of pixels operates with efficiency O (*N*) and has linear run time. Any mammogram image with *N* pixels requires a constant times *N* operations to cluster these pixels into groups.

The reason for linear efficiency is that for two pixels to be grouped into the same group, they must be nearest neighbors or be nearest neighbors of some combination of

pixels. This means that to cluster a single pixel, it is not necessary to consider every pixel in the image but only its eight nearest neighbor pixels. If a pixel belongs to a group with any other pixels, one pixel must be in its eight nearest neighbors. This allows processing of the linear list of pixels only once. Since the number of pixels in the image is much larger than eight, this is still considered order *N*.

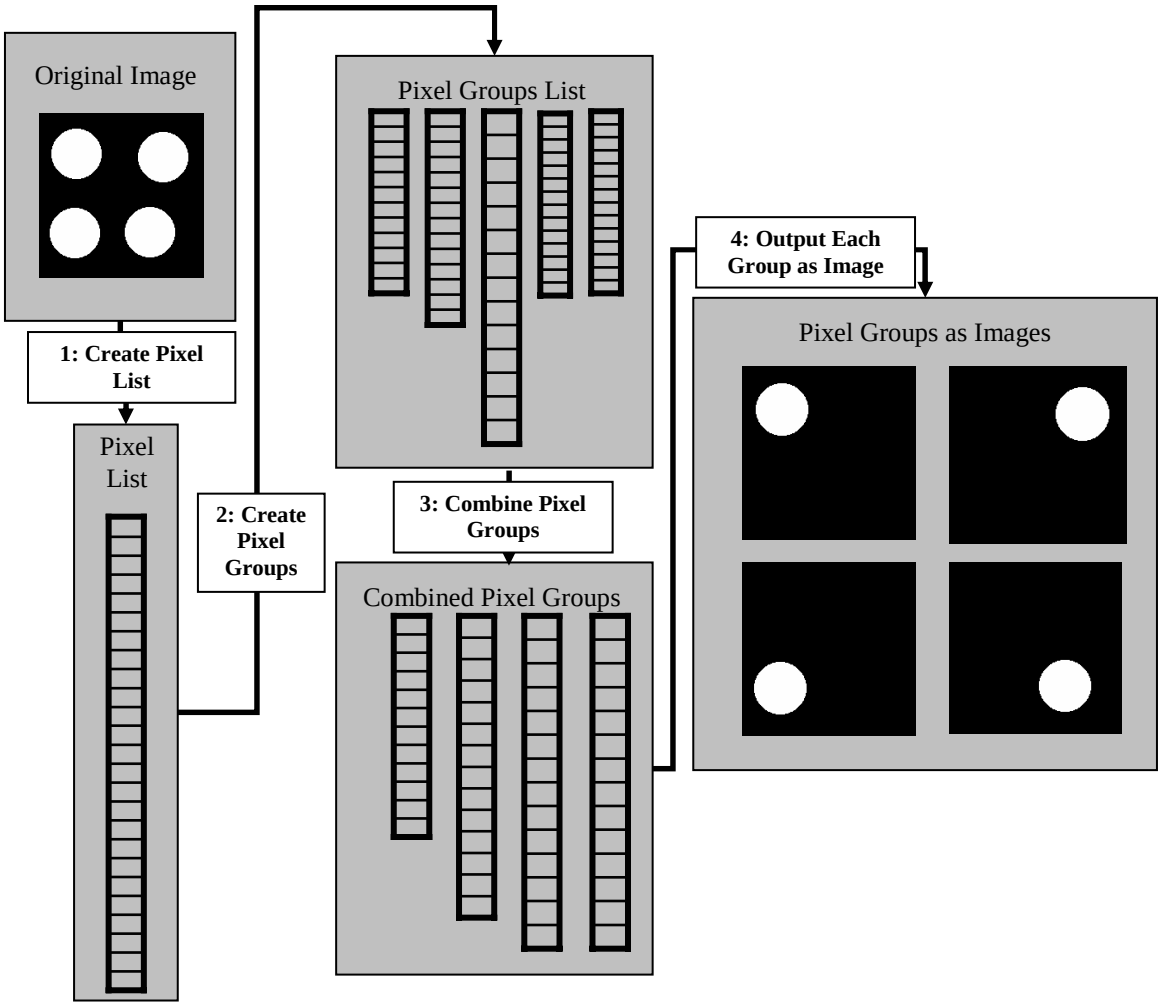


Figure 7.6: Diagram of pixel clustering algorithm

The initial processing of the pixel list produces the group-list. The length of this list is small compared to the number of pixels in the image. Step 3 of the algorithm combines the groups; however, some of these groups could be nearest neighbors. This is an effect of the simplification; if a pixel belongs to a group with any other pixels, one of

them must be in the eight nearest neighbors. Since the eight nearest neighbor pixels are the only pixels checked, there are instances in which a single group of pixels can be incorrectly classified in two groups. For example, Figure 7.7 contains only one distinct group of nearest neighbors. However, after two stages of the algorithm, two separate groups are in the group-list as shown on Figures 7.8 and 7.9.

The order the pixels are processed causes this; however, it is easily corrected. Each group in the group-list is checked and if they have common nearest neighbors, they are combined. This process is repeated until the groups are distinct. The process for combining the pixel groups that have nearest neighbors follows:

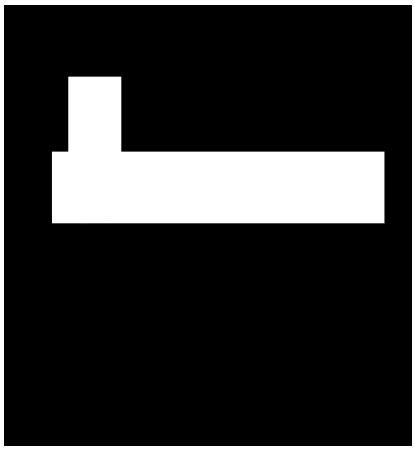


Figure 7.7: Example pixel cluster

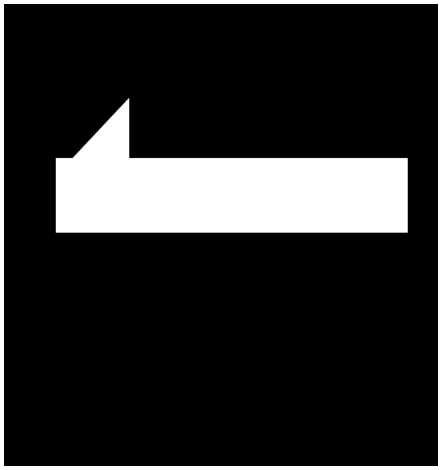


Figure 7.8: First incorrectly classified cluster

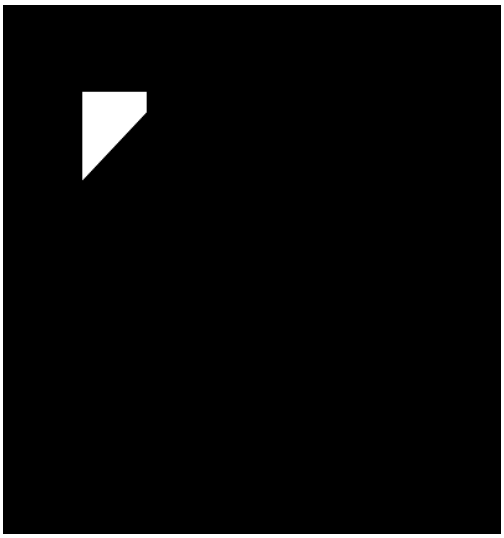


Figure 7.9: Second incorrectly classified cluster

Pixel Group Combination Algorithm

1. A new empty list of pixel groups is created.
2. The first from the original group list is added to this empty list.
3. The remaining groups are removed from the original list and added to the new list one at a time until the original list is empty. As each group is added to the new list, it is checked with the other groups. If it shares nearest neighbors with another group already on the list, it is combined with it. Once such a group is found the processing for that group ends and the next group begins.

Performed once, the possibility remains that groups in the new list still share nearest neighbors. This can occur when two groups are added to the new list, and a third group that connects the original two is added. It will be combined with only one of the original two. This possibility requires that the processing be repeated until the number of groups in the original list does not change. That is, no groups were combined and therefore no remaining groups share nearest neighbors.

This clustering algorithm can potentially require a substantial computation time, but in practice it performs a small number of iterations. If *G* represents the number of pixel groups after the initial processing of pixels, each step combining the groups can potentially require O (G^2) steps. The entire task of combining the groups can potentially require O (G^3) steps, and each group can potentially have *N/G* pixels. Using the simplification that each pixel is only checked against its immediate neighbors, each group comparison requires *N/G* steps. Therefore, the worst-case for the group combining algorithm is O (G^3) = O ((N/G)^3) = O (N^3). This would be unacceptable, but several conditions exist which mitigate this result. First, *G* is almost always very small compared to *N*. Since *G* represents significant areas of the mammogram image, it is usually less than fifty and the number of pixels in a mammogram image can be five million. When screening an image for mass detections, the number of pixels processed and the resulting number of pixel groups are often small (even zero). In addition, during the processing of the groups, when a group sharing nearest neighbors with another group is encountered, that step ends. That is, most steps in the process do not require the maximum number of computations.

In general, there is an inherent property in this algorithm, which prevents a worst-case scenario. If an image contains many small groups of pixels, the combining of the groups will be simplified by the small number of pixels in each group. When the image consists of a small number of large pixel groups, the small number of groups requires a minimum computation to process. By design, the use of this algorithm almost always satisfies one of these two conditions.

Two examples illustrate this, the first example shown on Figure 7.10 is an output image generated by our segmentation algorithm. Each shaded area represents a different

component of the mammogram image. By applying our pixel-clustering algorithm, these distinct components are extracted. These are shown as separate images on Figure 7.11; the background of Figure 7.10 is omitted.

Figure 7.10: Segmented image

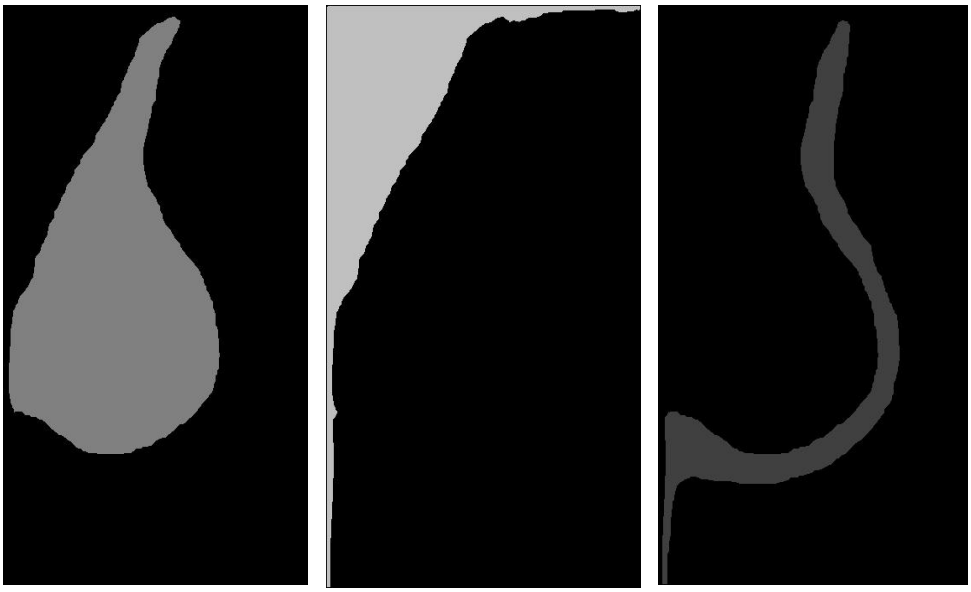


Figure 7.11: Image components

Each component in this example contains a large number of pixels. However, there are only four segments. Thus, the final step only requires O (*N*) steps. Since *G*=4, the combining step O (G^3) requires 64 group comparisons. In terms of *N*, each comparison step takes *N/G* steps, which is *N/*4. Thus the total group comparison step requires:

$$64 * \frac{N}{4} = 16 * N = O(N) \tag{7.1}$$

This is the efficiency of the first step in the clustering algorithm.

The second example is an output image from the mass detection algorithm, shown in Figure 7.12. It contains 20 distinct groups, and these groups contain an average of *N/*400 pixels. Thus the total group comparisons for this example require the following number of steps:

$$20^3 * \frac{N}{400} = 8000 * \frac{N}{400} = 20 * N = O(N) \tag{7.2}$$

This is also the efficiency of the first step in the clustering algorithm.

Once the detection results are clustered into distinct groups, they are placed in separate images and stored as files. This is a simple operation, but it permits a fast comparison with pathology.

In this dissertation research the clustered mass detections are compared directly with its pathology. However, one could use these detections to extract data from the original mammogram image for further processing and classification. This offers a significant increase in efficiency because the total area in these detected regions is much smaller than the entire image.

Figure 7.12: Mass detection results

7.3 Comparison with Pathology

Pathology accompanies each mammogram image used in this dissertation research. It is stored as a separate file, which details the abnormalities pathologists found in each image. The type of abnormality along with the image coordinates of a polygon that outlines the abnormality are contained in the pathology.

Using this polygon, a pathology image, which contains the location of the abnormality, is created. Figure 7.13 shows a mammogram image, the corresponding pathology image (which displays the location of a malignant mass) and the pathology image overlaid on the mammogram image.

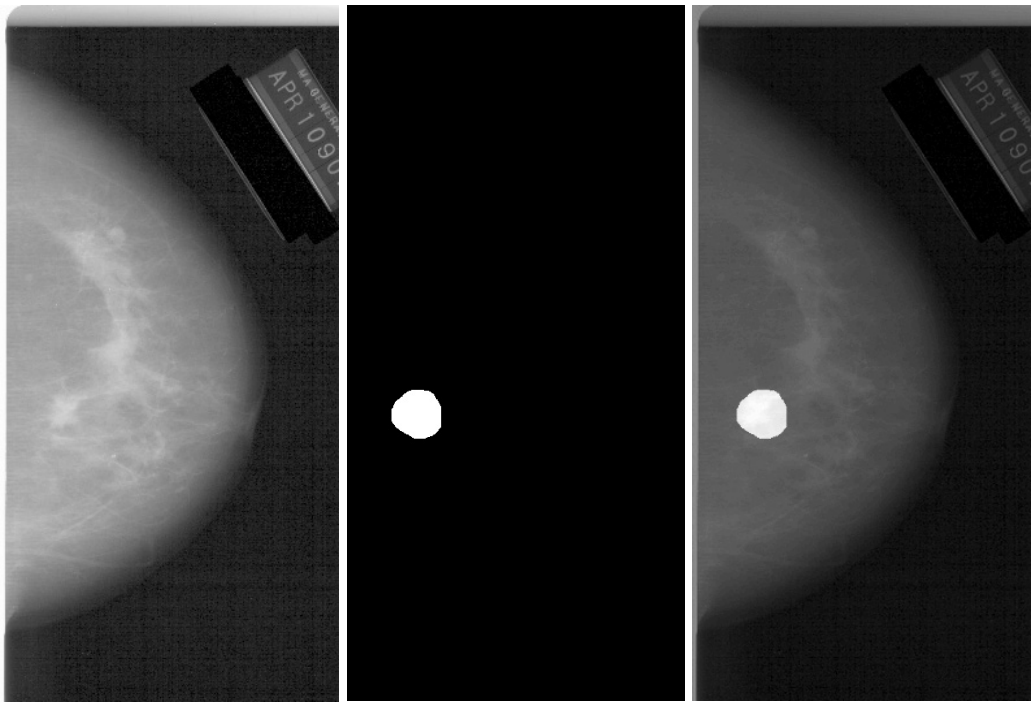


Figure 7.13: Mammogram image (left), pathology (center), and overlay (right)

Figure 7.14 shows the mass detection algorithm applied to the mammogram image in Figure 7.13. There are 4 groups of pixels shown on Figure 7.14, and one of these groups is the malignant mass indicated in the pathology. It is circled in the image. The other three probably represent erroneous detections. To determine this, the pixel-clustering algorithm is used to extract the four pixel groups from the image in Figure 7.14. Then each of these groups is compared to the pathology image in Figure 7.13.

To determine the accuracy of our mass detection algorithm, each of the detected masses is compared with the pathology image. We can determine if the areas intersect and by what percentage. This permits classification for each detection as true positive or false positive. True positives are detections that coincide with an actual malignant mass. False positives are detections that do not coincide with an actual mass. In addition, we can determine if a malignant mass was or was not detected by our algorithm.

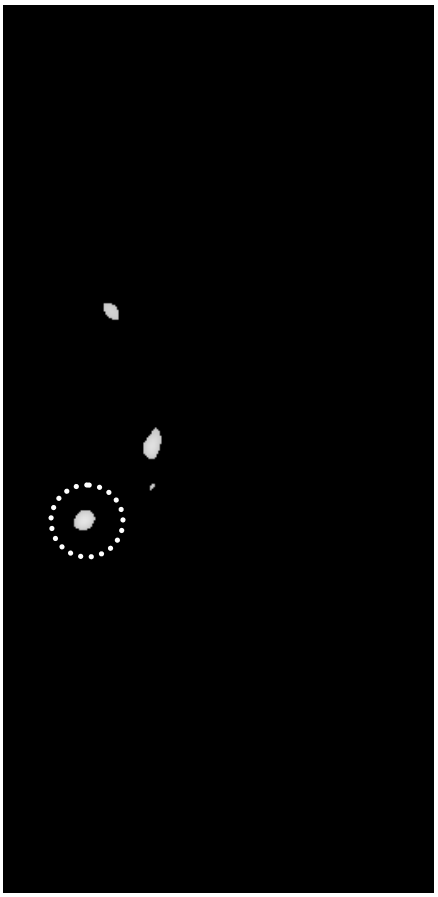


Figure 7.14: Results of mass detection algorithm

7.4 Process Parameters

A significant benefit of this system is the testing of the mass detection algorithm with a variety of parameters. In any image processing system development, there are many parameters to be chosen. Some parameters for the mass detection system are the mass template, the template size, and the threshold levels. To develop an effective method for locating masses, our system was tested with a large number of possible values for each of the parameters. For example, Figure 7.15 is the output of examining the mammogram image in Figure 7.1 with our mass detection algorithm. However, it has not been thresholded. Figure 7.16 shows the image after using thresholds of 0.70, 0.80, 0.85,

and 0.90. Figure 7.17 contains a table showing how the different thresholds affect the results of our algorithm.

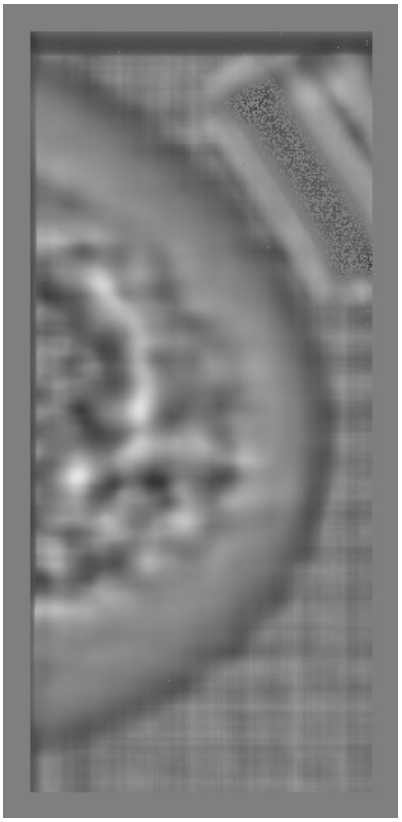


Figure 7.15: Results of mass detection algorithm

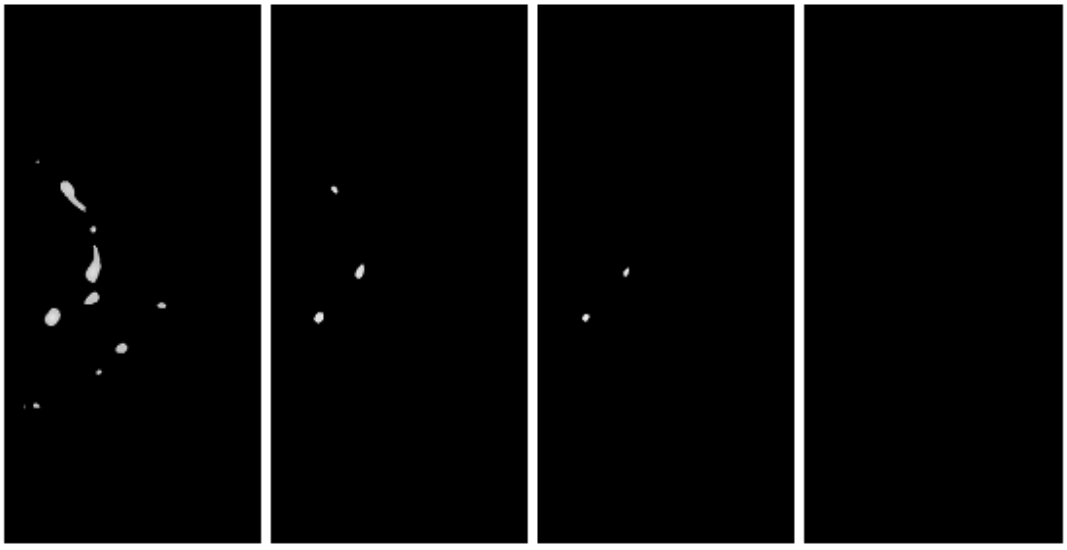


Figure 7.16: Results of thresholding with thresholds of 0.7, 0.8, 0.85 and 0.9

Table 7.1: Impact of image thresholds

Threshold Level	True Positives	False Positives
0.70	1	10
0.80	1	2
0.85	1	1
0.90	0	0

Clearly each threshold has a significant effect on the results. In this case, 0.85 is the optimal threshold setting. Our testing system allows us to quickly find that value and to also determine one set of parameters for all of the mammograms used.

7.5 Recording Results

The testing system records two statistics for each mammogram analyzed with the mass detection system. The first is the number of true positives. These are masses that were malignant and were found in the image. The statistic is a ratio of malignancies found to total malignancies. Obviously, the optimum result is a ratio of 1 to 1 or 0 to 0. Our second statistic is the number of false positives. This is the number of masses indicated by our technique that were not malignant. To classify an object as either a true positive or a false positive, the following definition for a true positive is offered. If an object returned by our technique occupies at least 50% of the area of an actual malignancy or the area of an actual malignancy occupies at least 50% of the object returned by our technique, it is considered a true positive. A false positive is an object returned by our technique which does not meet the area percentage for a true positive. Also, if a marked region contains more than one separate object, all but one are considered false positives. These results are reported for an entire set of mammogram images.

7.6 Process Summary

This brief summary of the testing system assumes that the output of the mass detection is stored in an image file with a predetermined naming convention.

The steps of the testing system are:

Testing System Process for a Single Mammogram Image

1. The output of the mass detection system, an image file, is read and thresholded according to an input parameter.
2. The pixel-clustering algorithm processes the thresholded image file and pixel clusters representing mass detections are stored as images according to the naming convention.
3. The pathology for the mammogram image is read and a "pathology" image is created.
4. The "pathology" image is compared to the mass detections. False positives and true positives are recorded.

Additions to the system allow this process to be automated. The system can perform over an entire set of images and for a preset variety of input parameters. Chapter 8 will present results from developing and testing the algorithms presented in this research. Those results are generated by this testing system and are used to generate plots of true positive sensitivity versus false positive errors. These plots are common for quantitatively comparing mass detection algorithms.

Chapter 8

Experimental Results

This chapter details the experimental methodology for testing the algorithms presented in this dissertation. It also presents the results of applying these algorithms to a set of mammogram images. The choice of this set of images for developing and testing and methods for reporting the performance of a diagnostic algorithm are discussed. In addition, how the testing system from Chapter 7 can be used to determine an optimal configuration is addressed. Detection results from multiple configurations of the mass location algorithm are also presented.

8.1 Developmental Image Sets

Several sets of mammogram images are available for public use in developing and testing mammogram-screening techniques. Publicly available images are very useful for this research for several reasons. First, taking images with x-ray film is costly and time-consuming. This would prohibit researchers without access to mammograms, film digitizers, and pathology from doing effective research. Also, images that are publicly available are open to widespread scrutiny for content, quality and usefulness. Only images that meet high standards will gain acceptance. Furthermore, the use of standard image sets adds credibility to results with those images. Researchers can compare their results with others' that used the same images.

For this research, the University of South Florida Digital Database for Screening Mammography (DDSM) containing over 10,000 mammogram images was used (Heath, Bower and Kopans). For our testing, a subset of this database called the Department of

Defense Breast Cancer Research Program image set was selected. This image set consists of four subsets: two for mass detection training and testing, and two for microcalcification training and testing. Different sets for training and testing are provided in this database. Researchers develop algorithms with the same training sets and then use the test sets for an unbiased verification of any results.

The mass training set has a total of 39 mammogram cases with 156 images and 80 tumors. The testing set has 40 mammogram cases with 160 images and 83 tumors. This provides efficient yet complete image coverage to develop and test our system. For this dissertation only the mass detection image sets were used because we do not address microcalcifications. The mass training set includes cases: 1118, 1134, 1156, 1159, 1160, 1163, 1166, 1174, 1203, 1212, 1217, 1222, 1224, 1229, 1236, 1252, 1262, 1403, 1417, 1467, 1486, 1520, 1557, 1587, 1589, 1592, 1620, 1622, 1642, 1671, 1693, 1700, 1701, 1720, 1726, 1790, 1896, 1899, and 1908. The testing set cases used are: 1112, 1114, 1122, 1127, 1140, 1147, 1149, 1155, 1168, 1169, 1171, 1207, 1211, 1228, 1233, 1234, 1237, 1247, 1258, 1401, 1416, 1468, 1485, 1504, 1510, 1573, 1577, 1618, 1628, 1658, 1669, 1673, 1674, 1804, 1821, 1827, 1892, 1906, 1985, and 1999.

8.2 Segmentation Algorithm Testing

8.2.1 Experimental Methodology

The segmentation algorithm presented in Chapter 5 used a histogram-based method to segment mammogram images into distinct components: breast composition, pectoral muscle, subcutaneous fat, and image background. We measured the performance of the segmentation algorithm in the following three ways.

1. All the segmented images are visually inspected to determine if the detected boundaries are reasonable, even though the boundaries are not always perfectly defined.
2. The pathology of each image is checked to verify and confirm that areas containing masses were not misclassified nor discarded. This is important because the segmentation process eliminates unnecessary areas from consideration when these areas are screened by the mass detection algorithm.
3. The total area of the image eliminated from consideration is determined. For this research background areas and pectoral regions are eliminated. Only the breast tissue is inspected for signs of abnormalities. However, some masses may be projected into the pectoral region in the image. In these instances, the mass location algorithm analyzes the pectoral region separately.

8.2.2 Results

The segmentation algorithm was tested with 156 images from the DDSM training image set. The images contained a total of 80 tumors identified by radiologists and confirmed by pathologist with the pathology included. When the segmentation algorithm is applied to these images, 46% of the pixels in the images are eliminated from further consideration. This is determined by counting the total number of pixels in the 156 images, and then comparing this total with the number of pixels classified as in the image background or pectoral muscle. Table 8.1 presents the results of applying the mass detection algorithm in Chapter 6 including and not including the preprocessing segmentation. Analysis of the segmented images produces nearly identical results with a slight reduction in the number of false positives per image compared to analysis of non-

segmented images. Thus, our method offers a significant improvement in computational efficiency with no degradation in the diagnostic accuracy. Also, none of the masses in these test images were classified as background or pectoral muscle, thus none were eliminated from further consideration.

Table 8.1: Detection results including and not including image segmentation

	False Positives Per Image	
Accuracy	With Segmentation	Without Segmentation
0.9625	56.13	55.88
0.9625	33.69	34.08
0.9625	16.34	16.51
0.7625	4.46	5.63
0.3750	0.97	1.05

8.3 Testing of Mass Detection Algorithm

8.3.1 Experimental Methodology

The mass location algorithm presented in Chapter 6 was developed using images from the training set from the DDSM. The examination of mass templates was also done with these images. Once the method was complete, it was verified with images from the DDSM testing set. The testing images are not used in the development of the algorithm, so the results are not biased.

To measure the performance of the mass location algorithm, the testing system described in Chapter 7 are used. This system extracts detected masses from screening results and compares these detections with the pathology. This testing system is designed to accommodate multiple combinations of parameters that are used to refine and optimize the algorithm. The results presented in the following sections present the effects of varying those parameters. Results are presented on free-response receiver operating

characteristic (FROC) curves and in tables that show the performance relationships between the various parameters and the screening algorithm.

8.3.2 Free-Response Receiver Operating Characteristic Curve

The most used method for presenting diagnostic performance of computer aided screening algorithms is the use of a free-response receiver operating characteristic (FROC) curve, a variant of a receiver-operating characteristic (ROC). These were initially designed to evaluate detectability in radar (Beutel, Kundel and Van Metter 752), but have commonly been adapted to medical decision-making. The FROC curve plots diagnostic sensitivity versus false positives. Specifically, the vertical axis of the plot represents the percentage of masses that were correctly found for a specified sensitivity. The horizontal axis represents the number of false positives per image. Thus, the curve shows the relationship between diagnostic sensitivity and false positives. The goal of almost all screening algorithms is to maximize the percentage of masses found, while minimizing the number of false positives per image.

FROC plots allow researchers to examine the properties of their algorithms. For example, when the mass location algorithm is applied, a new image is returned which indicates the areas likely to contain a mass. We can threshold that image with any value between 0 and 255. A threshold of 0 would eliminate no areas, and return 100% of the masses in the test set. However, it would also return an unacceptably high number of false positives. A threshold of 255 would eliminate all areas of the image and return no false positives. Of course, it would also not locate any masses in the test set. For a method to be useful and effective, a proper threshold must be determined. The

relationship between the threshold and diagnostic sensitivity, from FROC curves, is used to refine and help determine the proper threshold.

8.3.3 Results for Initial Mass Location

These following results were gathered by examining the 160 images in the test image set, which contains a total of 83 malignant masses. The first set of results is obtained with the mass detection algorithm using a single template on each image in the test set. These results are thresholded, clusters of pixels are extracted, and the clusters are compared to the pathology.

These first results demonstrate the effect of varying the template size and the threshold. Table 8.2 shows the results for 25×25 pixel template, which after scaling is actually 0.4mm×0.4mm. In this first set of results, we can view the effect of varying the threshold in the image 180-to-210. Note that the correlation results vary from -1 to 1, and are scaled to the integer range of 0 to 255. This allows these results to be stored like images. When the results are thresholded with the value 180, 100% of the masses are located (found). This is a significant result because it shows that this screening system is extremely effective at locating masses. However, with 100% detection performance (sensitivity), there are 36.66 false positives per image, a higher number than desired. As the threshold value is increased from 180 to 210, the number of false positives per image decreases from 36.66 to 1.18. 1.18 false positives per image is a very acceptable number, however, the detection performance is 38.6%. The FROC curve in Figure 8.1 displays the relationship between diagnostic sensitivity and false positives per image. (The perfect result for a system is 0 false positives when all masses are detected.)

Next, the same process is repeated but the template size is changed to 50×50. The results are shown on Table 8.3. Figure 8.2 is the FROC curve for this. The change in template size from 25×25 to 50×50 pixels has a marked effect on results and demonstrates the importance of testing the system with various parameters.

Table 8.2: Mass location results for template size 25×25

Threshold	False Positives Per Image	True Positives	Missed Masses	Sensitivity
180	36.66	83	0	100%
185	27.11	81	2	97.6%
190	18.58	79	4	95.2%
195	11.43	77	6	92.8%
200	6.23	70	13	84.3%
210	1.18	32	51	38.6%

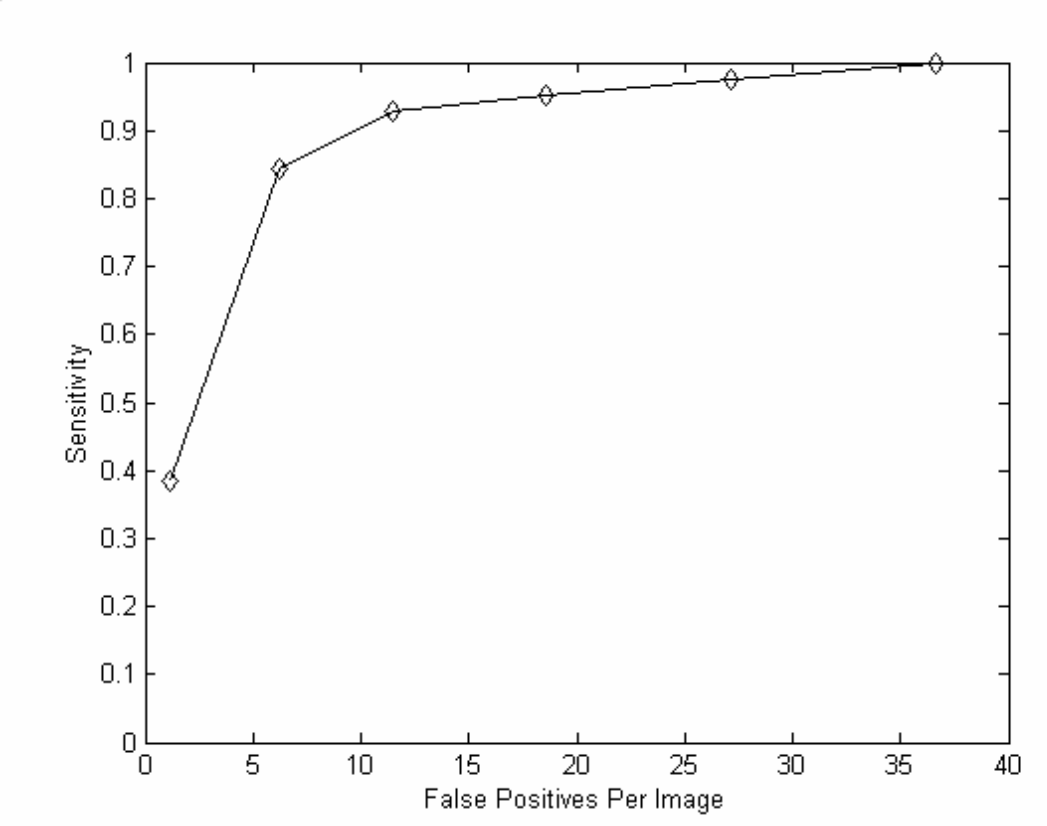


Figure 8.1: FROC curve of mass location results for template size 25×25

Table 8.3: Mass location results for template size 50×50

Threshold	False Positives Per Image	True Positives	Missed Masses	Accuracy
180	12.44	76	7	91.6%
185	9.63	74	9	89.2%
190	7.09	72	11	86.8%
195	5.07	71	12	85.5%
200	3.36	63	20	75.9%
210	1.24	44	39	53.0%

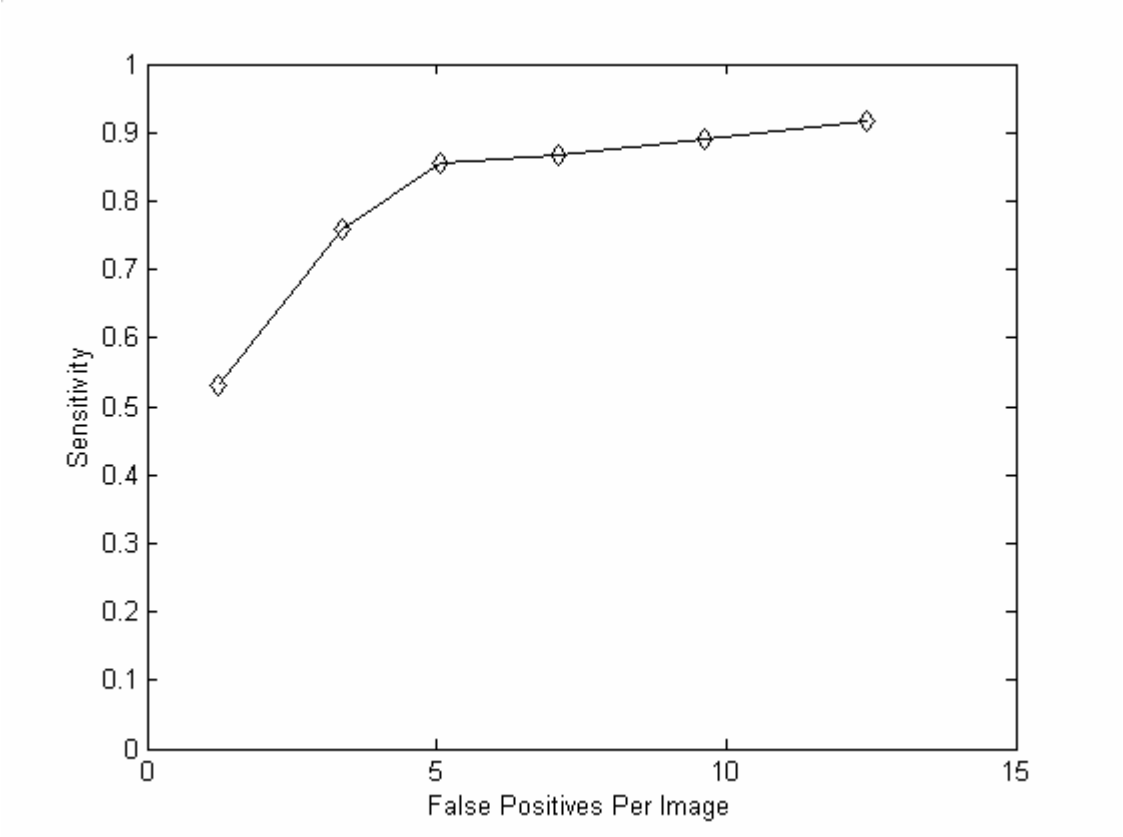


Figure 8.2: FROC curve of mass location results for template size 50×50

When the results for mass location using the 50×50 pixel template were thresholded with 180, the sensitivity is 91.6% with 12.44 false positives per image. This is similar to the results from the previous example with a 195-threshold level. In fact, by varying the threshold levels, we can produce similar results with mass templates of different sizes. However, using a 50×50 template requires four times longer to compute

than the 25X25 template does. Thus, we want to choose the smallest template size with acceptable performance.

The images were tested with template sizes smaller than 25×25, but their results were very poor. Templates smaller than 20×20 lose the characteristic shape that makes them effective. From the initial results, we determined that the 25×25 pixel template produces the best results. Thus, the following results were produced with an initial template size of 25×25 pixels.

8.3.4 Results for Multiscale Mass Location

To reduce the number of false positives returned by the mass location algorithm, a multiscale step described in section 6.6 was developed. The multiscale step uses the potential mass detections and rescales one or more templates to the size and location of each potential mass. Then, the mass detections and new templates are compared using statistical correlation and the results are thresholded with a predetermined sensitivity.

We know if the template is the proper scale and centered exactly on the mass, then correlation will be close to 1. For the 80 masses in the training set, the average correlation between the mass and the scaled-centered template was 0.8102 with a standard deviation of 0.0476. The correlation values for each of the 80 masses are shown in Figure 8.3. The histogram of these values is shown in Figure 8.4. The histogram shows a fairly tight distribution. If the threshold were set as high as 0.7, only 3 tumors would be eliminated. This allows the use of a multiscale step to help reduce the number of false positives returned by the mass location algorithm.

To apply the multiscale step, the extracted mass detections returned from the example with the 25×25 pixel templates were collected. Since we are trying to eliminate

false positives in this second step, the most inclusive results from the first step were used. These are the results from a threshold of 180, which found 100% of the masses in the set with 36.66 false positives per image. The mass detections from this first step are then compared with templates. The templates are scaled to 200%, 100%, 66% and 33% the size of the masses that are detected. The maximum correlation between the detected mass and the four scaled templates is thresholded. Mass detections above the threshold are considered final, and those below the threshold are discarded. The results for this are presented on Table 8.4 and its FROC curve is shown in Figure 8.5.

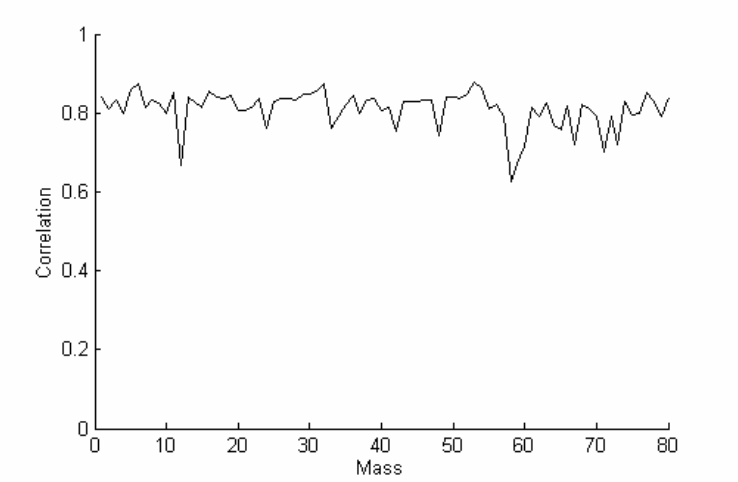


Figure 8.3: Correlation of scaled template for each mass in developmental set

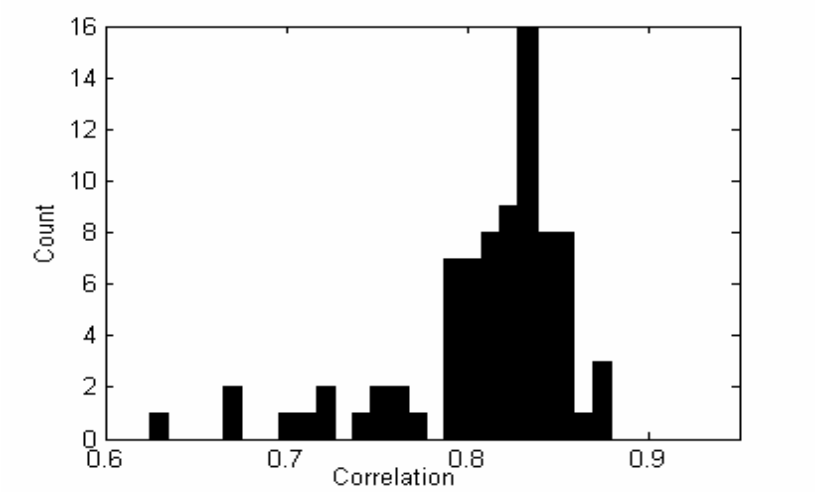


Figure 8.4: Histogram of mass to scaled template correlation

Table 8.4: Results for multiscale post processing

Threshold	False Positives Per Image	True Positives	Missed Masses	Accuracy
170	36.20	83	0	100%
180	34.91	83	0	100%
185	32.69	83	0	100%
190	29.18	81	2	97.59%
195	24.22	81	2	97.59%
200	19.19	81	2	97.59%
210	11.51	69	13	84.33%

This post-processing step helps reduce the number of false positives returned. Before multiscale post processing, the mass detection algorithm had 36.66 false positives per image with 100% sensitivity, and with the post processing the number of false positives are reduced to 32.69 per image.

8.3.5 Results with Multiscale Templates at Initial Examination

In the previous section, a method to add multiscale functionality to the process by reexamining potential mass detections with a template scaled to match the size of the potential mass detection was described. Another possibility to test the data with multiple templates is to do the initial template-to-image correlation for several different size templates in the same pass. We can either average the correlations from multiple templates or use the maximum (best) correlation with each mass in the original image.

We use the 25×25 pixel template for an initial examination with scales of 100%, 66%, and 33%. In the first example, the correlations from the three scales are averaged and used as the resultant image. In the second example, the maximum (best correlation) of the three scales is used as the resultant image. Results from the averages example are shown in Table 8.5 and Figure 8.6. Results from the maximum (best correlation) example

are shown in Table 8.6 and Figure 8.7. The results from both are similar to the single scale approach and do not produce significant improvement to the results. This agrees with a similar result by Brake and Karssemeijer.

8.3.6 Results with Limiting the Number of Detections

One method for reducing the number of false positives returned by a mass location method is to limit the number of allowed detections. This is accomplished by ranking all the detections for each image and accepting a maximum fixed number of detections. In this, the template size is 25×25 pixels and each potential mass is ranked by the average correlation of the pixels in that mass. A test with the maximum number of masses at 5, 10, 20, 30, and 40 is done. Table 8.7 presents these results.

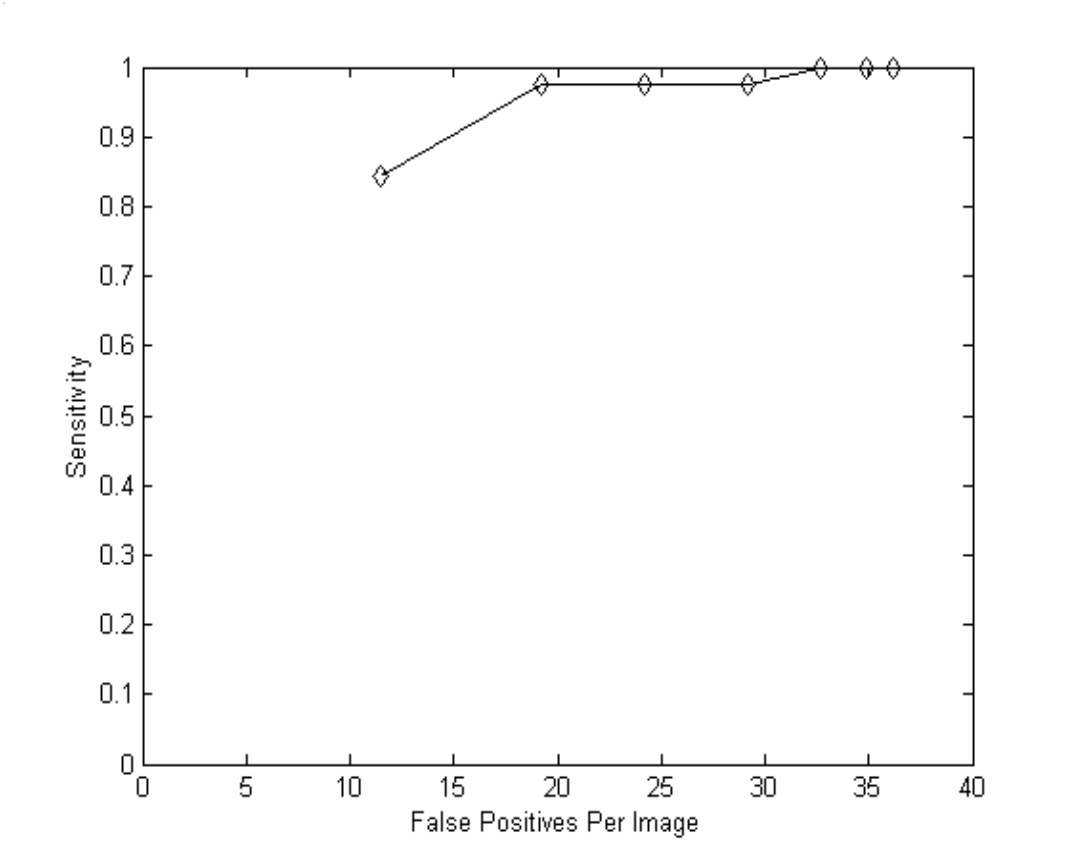


Figure 8.5: FROC curve for multiscale post processing

Table 8.5: Results from averaging multiple template scales for initial examination

Threshold	False Positives Per Image	True Positives	Missed Masses	Accuracy
180	11.55	77	6	92.77%
185	7.76	74	9	89.16%
190	4.89	71	12	85.16%
195	2.84	61	22	73.49%
200	1.48	46	37	55.42%
210	0.22	20	63	24.10%

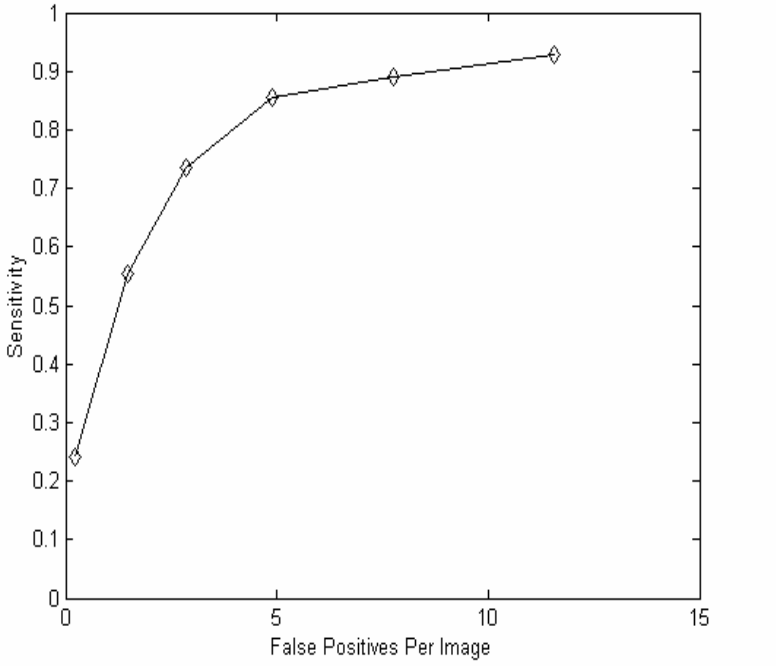


Figure 8.6: FROC curve for averaged multiscale examination

Table 8.6: Results from selecting maximum correlation of multiple template scales for initial examination

Threshold	False Positives Per Image	True Positives	Missed Masses	Accuracy
180	33.41	82	1	98.80%
185	26.72	80	3	96.39%
190	19.88	78	5	93.98%
195	13.74	79	4	95.18%
200	8.77	79	4	95.18%
210	2.68	66	17	79.52%

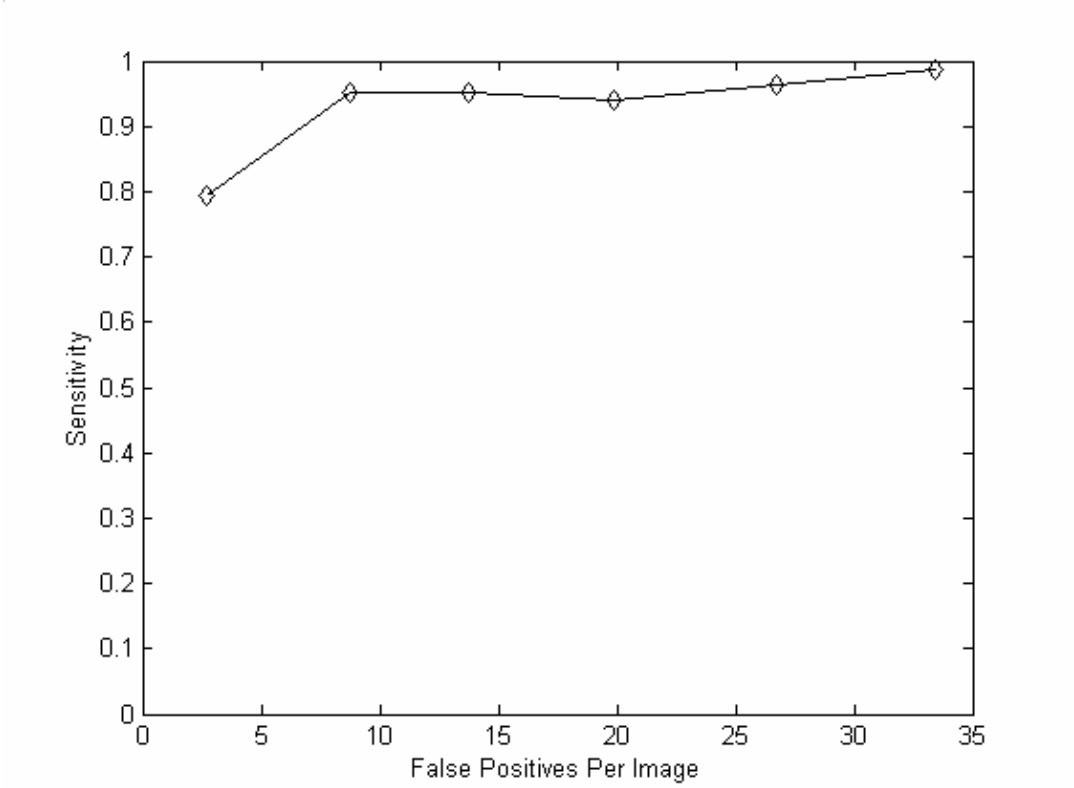


Figure 8.7: FROC curve for maximum correlation example

Table 8.7: Results with limits to number of detections

Detections Limit	Threshold Value	False Positives Per Image	Sensitivity	Percent of False Positives Reduced	Sensitivity Reduction
5	180	4.8187	34.94%	86.86%	65.06%
10	180	9.7562	46.99%	73.39%	53.01%
20	180	19.6125	63.86%	46.50%	36.14%
30	180	33.2750	93.98%	9.23%	6.02%
40	180	36.5812	100%	0.21%	0%

This produces a reduction in false positives, but it also reduces the number of true detections. In fact, the reduction in false positives is only slightly higher than the reduction of true detections. Therefore, this step does not improve the performance of our algorithm. However, it does show that our system's detections are almost evenly distributed within the false positives according to this ranking system for possible mass

detections. Since the mass location and ranking are both template matching based, we can conclude that the limit has been reached for possible results with only the mass template. To further reduce the number of false positives, a new-and-different ranking system is required.

8.4 Discussion

The results and examples presented in this chapter are a thorough review of both the mammogram segmentation and mass location methods. For the mass detection algorithm an optimal configuration was determined. The mass location system located 100% of the masses with one configuration but returned 36.66 false positives per image in the set. The sensitivity of this algorithm is very significant. This number of 36.66 false positives per image is too high for presentation to a radiologist. Several methods to reduce the number of false positives were presented, but none were sufficiently effective. Thus, the mass location system presented in this research is best used as a screening algorithm because it is very effective at locating potential masses in mammograms. To reduce the number of false positives, another algorithm needs to be developed which determines differences between malignant and benign masses.

Chapter 9

Summary

9.1 Review

A new-unique algorithm for segmenting mammogram images into the image background, breast composition, pectoral muscle, and subcutaneous fat is presented. Results and examples of the outstanding performance for this algorithm were presented. This segmentation serves as a pre-processing step in the screening process. It significantly reduces the number of pixels that need to be used in the mass-screening algorithm. This segmentation method permits each component to be examined independently of the others with additional searches, which can reduce errors and simplify the search. It has the potential for use by existing adaptive detection algorithms.

A new method is also presented to locate mass lesions in a mammogram image. The technique uses template matching to locate a potential mass in the image and applies a multiscale approach to reduce the number of false positives. The template-matching algorithm was evaluated with several different potential templates. This method can also be used as a pre-processing step. When desirable, it can extract suspicious areas for further analysis. It can focus and limit the areas to be searched for masses in an image, which will reduce the search time in a more complex analysis.

A system was developed for the general testing of screening algorithms, recording their results, and comparing those results with pathology. It makes the testing of many multiple configurations in future research into mass lesion location feasible. Additional

configurations and parameters were studied with this system and these results were presented.

9.2 Future Research

The research in this dissertation significantly improves results for some of the most important problems that remain in computer assisted screening of digital mammogram images. Before these new algorithms are used in commercial screening, one important research problem remains.The remaining research would focus on a more accurate way to distinguish between malignant and benign masses. The mass location method, presented in this dissertation, finds essentially all of the known malignant masses, which is excellent; however, it also detects benign masses and other non-cancerous mass-like objects superbly. The total number of benign detections is too large and produces too many prompts. A maximum of 4 to 5 prompts per image may be acceptable but no more than 2 per image would definitely be acceptable. Our algorithm sometimes returns over 32 detections per image (most of which are benign) with 100% sensitivity and over 11 detections per image with 92.8% sensitivity. These results represent the best performance with our current template.

9.3 Conclusion

The algorithms developed in this dissertation provide for an excellent-automatic mass screening for mammogram images and for segmenting the mammogram image. The automatically screening for mass lesions is an important contribution. It is also significant that this mass location system is very effective in locating mass lesions while still being computationally efficient.

References

American Cancer Society. Breast Cancer Death Rates Continue to Decline. 10 Oct. 2001. <http://www.cancer.org/docroot/NWS/content/NWS_1_1x_Breast_Cancer_Death_Rates_Continue_to_Decline.asp>

Beutel, Jacob, Harold L. Kundel and Richard L. Van Metter. Handbook of Medical Imaging Volume 1. Physics and Psychophysics. Bellingham, Washington: SPIE Press, 2000.

Brake, Guido M.te, and Nico Karssemeijer. "Single and Multi-scale Detection of Masses in Digital Mammograms." IEEE Transactions on Medical Imaging 18:7 (1999):628-638.

Brzakovic, D., et al. "An Approach to Automated Detection of Tumors in Mammograms." IEEE Transactions on Medical Imaging 9:3 (1990): 233-241.

Cady, Blake and James S. Michaelson. "The Life-Sparing Potential of Mammographic Screening" Cancer 91:9 (2001): 1699-1703.

Chan, H-P, et al. "Improvement in Radiologists' Detection of Clustered Microcalcifications on Mammograms: The Potential of Computer aided Diagnosis." Investigative Radiology 23 (1990): 664-671.

Dhawan, Atam P., et al. "Analysis of Mammographic Microcalcifications Using Gray-Level Image Structure Features." IEEE Transactions on Medical Imaging 15:3 (1996): 246-59.

D'Orsi, Carl J. et al. Illustrated Breast Imaging Reporting and Data System. Reston: American College of Radiology, 1998.

Gale, A.C. and B.S. Worthington "Visual Attention in Diagnostic Radiology," Contemporary Ergonomics. Ed. D.J. Oborne. Burlington, Taylor and Francis, 1986.

Gonzales, R.C. and Paul Wintz. Digital Image Processing. Reading, MA: Addison-Wesley, 1977.

Heath, M., K.W. Bowyer and D. Kopans, "Current status of the Digital Database for Screening Mammography," Digital Mammography 1998:457-60.

Hutt, Ian. "Computer aided Detection of Abnormalities in Digital Mammograms." Diss. U. of Manchester, 1996.

Karssemeijer, Nico, J.T.M Frieling and J.H.C.L Hendricks "Spatial Resolution in Digital Mammography." Investigative Radiology 28 (1993): 413-9.

Kim, Jong Kook and Hyun Wook Park. "Statistical Textural Features for Detection of Microcalcifications in Digitized Mammograms." IEE Transactions on Medical Imaging 18.3 (1999): 231-8.

Kobatake, Hidefumi, et al. "Computerized Detection of Malignant Tumors on Digital Mammograms." IEEE Transactions on Medical Imaging 18:5 (1999): 369-78.

Kundel, H.L. and C.F. Nodine. "Studies of Eye Movements and Visual Search in Radiology." Eye Movements and the Higher Psychological Functions. Ed. J.A.W. Seder. et al., Hillsdale, NJ: LEA, 1978.

Kupinski, Matthew A. and Maryellen L. Giger. "Automated Seeded Lesion Segmentation on Digital Mammograms." IEEE Transactions on Medical Imaging 17:4 (1998): 510-517.

Li, H.D., et al., "Markov Random Field for Tumor Detection in Digital Mammography", IEEE Transactions on Medical Imaging 14:3 (1995): 565-576.

Li, Huai, et al. "Computerized Radiographic Mass Detection—Part I: Lesion Site Selection by Morphological Enhancement and Contextual Segmentation." IEEE Transactions on Medical Imaging 20:4 (2001): 289-301.

Linnett, L.M., et al. "Texture Classification Using a Spatial-Point Process Model." IEEE Proc. Vis. Image Signal Process 142:1 (1995): 1-6.

Morrison, Steven and Laurie M. Linnett. "A Model Based Approach to Object Detection in Digital Mammography." Proceedings 1999 International Conference on Image Processing (1999): 182-6.

Petrick, Nicholas, et al. "An Adaptive Density-Weighted Contrast Enhancement Filter for Mammographic Breast Mass Detection", IEEE Transactions on Medical Imaging 15:1 (1996): 59-67.

Sivaramakrishna, R. et al. "Automatic Segmentation of Mammographic Density", Academic Radiology 8:3 (2001)

Stamatakis, E.A. et al. "Detecting Abnormalities on Mammograms by Bilateral Comparison", Proceedings of IEE Colloquium on Digital Mammography London: IEE, 1996.

Strickland, Robin N. and Hee Il Hahn. "Wavelet Transforms for Detecting Microcalcifications in Mammograms." IEEE Transactions on Medical Imaging 15:2 (1996): 218-229.

Sucking J., et al. “Segmentation of Mammograms Using Multiple Self-Organizing Neural Networks” Medical Physics 22:2 (1995): 145-152.

United States. National Institutes of Health. What you need to know about breast cancer. NIH publication no. 00-1556. Washington: National Cancer Institute, 2000 <http://www.cancer.gov/cancerinfo/wyntk/breast>.

United States. National Institute of Statistics. NIST/SEMATECH e-Handbook of Statistical Methods. 2002 <http://www.itl.nist.gov/div898/handbook/>

University of California, Davis Health System. What is Breast Cancer?. March 2002. <http://www.ucdmc.ucdavis.edu/ucdhs/health/a-z/06BreastCancer/doc06full.html>

Veldkamp, Wouter J.H. and Nico Karssemeijer. “Normalization of Local Contrast in Mammograms.” IEEE Transactions on Medical Imaging 19:7 (2000): 731-8.

Yu, Songyang and Ling Guan. “A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films.” IEEE Transactions on Medical Imaging 19:2 (2000): 115-126.

Vita

John Sample was born on January 12, 1977, in Texarkana, Texas. He received his bachelor's degree in mathematics from the University of Southern Mississippi in 1999. In August 1999, he began graduate studies in computer science at Louisiana State University as a Board of Regents Fellowship recipient. His research interests include image processing, parallel computing, and geospatial information systems.